# Multimodal E-assessment:

# An Empirical Study

**Amirah Nasser Algahtani**

**Thesis submitted for the degree of Doctor of Philosophy**

**in E-learning**

**Supervised by: professor Dimitrios Rigas**

**Faculty of Technology, De Montfort University**

**2015**

# AN EMPIRICAL INVESTIGATION IN MULTIMODAL E-ASSESSMENT

## The effect of multimodal metaphors on the usability and communication performance of e-assessment

**Amirah Nasser Algahtani**

**Thesis submitted for the degree of**

**Doctor of Philosophy in**

**E-learning**

**Supervised by:**

**Professor Dimitrios Rigas**

**Faculty of Technology, De Montfort University**

# List of Publications

**Journal Contributions**

- D. Rigas and A. Algahtani." Multimodal E-assessment Interfaces: An Empirical Investigation." Europment 2015- Journal, Vol 978-1-61804-254-5

**Conference Contributions**

- A. Algahtani and D. Rigas.“ Multimodal E-assessment Interfaces: An Empirical Investigation" Human-Computer Interaction. World Scientific and Engineering Academy and Society, 2015.
- D. Rigas and A. Algahtani. "The role of expressive, full-body avatars and earcons and auditory icons in e-assessment interfaces." Europment 2014- Athens, ISBN: 978-1-61804-254-5
- D. Rigas and A. Algahtani. "An Investigation of Multimodal Metaphors in E-assessment Interfaces." HCI International 2013-Posters’ Extended Abstracts. Springer, Berlin Heidelberg, 2013. 567-571.
- A. Algahtani and D. Rigas.“ On-Line Assessment: An Audio-Visual Approach" Human-Computer Interaction. World Scientific and Engineering Academy and Society, 2013.

# LIST OF ACRONYMS

| | |
|---|---|
| VMAP | Multimodal E-assessment condition |
| VOAP | Without Multimodal E-assessment condition |
| CV | Controlled Variable |
| df | degree of freedom |
| DV | Dependent Variable |
| E-assessment | Electronic- assessment |
| AI | Avatar with images |
| RI | Record with images |
| DI | Description text with images |
| ABGC | Avatar Body Gestures condition |
| H | Hypothesis |
| HCI | Human Computer Interaction |
| ICT | Information and Communication Technology |
| IV# | Independent Variable |
| E-Learning | Electronic Learning |
| MD | Mean Definition |
| SUS | System Usability Scale |
| ANOVA | Analysis Of Variance |
| VB | Visual Basic |

# DEDICATION

This thesis is dedicated to my husband Badr Almutairi the twin of my soul, and our dear son Mohammed, and daughters Wasan, Lana and Seema for their constant love, encouragement and patience.

# ABSTRACT

Due to the availability of technology, there has been a shift from traditional assessment methods to e-assessment methods designed to support learning. With this development there is a need to address the suitability and effectiveness of the e-assessment interface. One development in the e-assessment interface has been the use of the multimodal metaphor. Unfortunately, the associated effectiveness of multimodality in terms of usability and its suitability in achieving assessment aims has not been fully addressed. Thus, there is a need to determine the impact of multimodality on the effectiveness of e-assessment and to reveal the benefits, primarily to the user. Moreover, those involved in the development and assessment should be aware of potential impacts and benefits. This thesis investigates the role and effectiveness of multimodal metaphors in e-assessment, specifically; the thesis assesses the effect of multimodal metaphors, alone or in combination, on usability in e-assessment. Usability includes efficiency, effectiveness and user satisfaction. The empirical research described in this study consisted of three experiments of 30 participants each to evaluate the effect of description text, avatars and images individually, avatars, description text and recorded speech in combination with images, and finally, the use of avatars with whole body gestures, earcons and auditory icons. The experimental stages were designed as a progression towards the main focus of the study, which was the effectiveness of full body gesture avatar, considered to be the latest development in multimodal metaphors. The experimentation also assessed the role that an avatar could play as a tutor in e-assessment interfaces. The results proved the positive effectiveness and applicability of metaphors to enhance e-assessment usability. This was achieved through a more effective interaction between the user and the assessment interface. A set of empirically derived guidelines for the design and use of these metaphors to enhance e-assessment is also used in order to generate more usable e-assessment interfaces.

## Table of Content


**List of Publications** .......... II

**List of Acronyms** .......... III

**Dedication** .......... IV

**Abstract** .......... V

**List of Figures** .......... XI

**List of Tables** .......... XII

**Chapter 1** .......... 1

**Introduction** .......... 1

- 1.1 Motivation .......... 1
- 1.2 Aims .......... 4
- 1.3 Objectives .......... 4
- 1.4 Overall Hypothesis .......... 5
- 1.5 Methodology .......... 5
- 1.6 Thesis Contribution .......... 7
- 1.7 Outline .......... 8

**Chapter 2** .......... 10

- Multimodality and E-assessment Interface .......... 10
- 2.1 Introduction .......... 10
- 2.2 E-assessment .......... 11
  - 2.2.1 Definition of E-assessment .......... 12
  - 2.2.2 Benefits .......... 12
  - 2.2.3 Limitations .......... 13
  - 2.2.4 Challenges .......... 14
  - 2.2.5 Usability Evaluation in E-assessment Interface .......... 16
- 2.3 Multimodality .......... 18
- 2.4 Visual Display .......... 21
  - 2.4.1 Auditory Modal .......... 22
  - 2.4.2 Speech .......... 22
  - 2.4.3 Recorded Speech .......... 23

2.4.4 Non-speech .......... 23
2.4.5 Auditory Icons .......... 24
2.4.6 Earcons .......... 25
2.5 Avatars .......... 26
2.5.1 Body Gestures .......... 27
2.6 Multimodal Systems Interaction .......... 28
2.7 Why Multimodality .......... 28
2.8 Humanising Interfaces .......... 29
2.9 Multimodal in E-assessment Critical Review .......... 30
2.10 Summary .......... 31
Chapter 3 .......... 32
Experimental Phase I: The Role of Multimodal E-assessment Interfaces .......... 32
3.1 Introduction .......... 32
3.2 Aims .......... 32
3.3 Hypothesis .......... 33
3.4 Design of Experimental Condition .......... 34
3.4.1 Non Multimodal Condition (VOAP) .......... 35
3.4.2 Multimodal E-assessment Condition (VMAP) .......... 35
3.4.3 Implementation of Avatars .......... 38
3.4.4 Implementation of Images .......... 38
3.4.5 Implementation of Description Text .......... 38
3.4.6 Design of Tasks .......... 39
3.5 Variables .......... 40
3.5.1 Independent Variables .......... 40
3.5.2 Dependent Variable .......... 40
3.5.3 Controlled Variable .......... 40
3.6 Data Gathering and Questionnaire .......... 41
3.7 Pilot Study .......... 42
3.8 Users Sampling .......... 42
3.9 Results and Analysis .......... 44
3.9.1 Users Profiling .......... 44
3.9.2 Experiment Sessions .......... 45
3.9.3. Individual User .......... 46

3.9.4 Efficiency ........................................................................................................46
3.9.5 Effectiveness ..........................................................................................................49
3.10 Satisfactions ........................................................................................................ 52
3.11 Discussion........................................................................................................... 53
3.12 Summary ............................................................................................................ 56
Chapter 4 .................................................................................................................58
Experiment II: The Role of Avatar in E-assessment............................................................ 58
4.1 Introduction ........................................................................................................ 58
4.2 Aims .................................................................................................................... 58
4.3 Tasks.................................................................................................................... 59
4.4 Hypotheses ......................................................................................................... 60
4. 5 Experimental Condition..................................................................................... 60
4. 5. 1 First Interface Design (Record Speech and Images RI) ..........................................61
4. 5. 2 Second Interface Design (Description Text and Images DI) ...................................63
4.5. 3 Third Interface Design (Avatar and Images AI).......................................................63
4. 5. 4 Experiments Design ................................................................................................64
4.5.5 Experiment Tasks ......................................................................................................64
4.5.6 Implementation of Avatars and Record Speech and Description Text....................................65
4.5.7 Pre-Experimental Questions.......................................................................................66
4.6 Variables ............................................................................................................. 66
4.6.1 Dependent Variables ...............................................................................................66
4.6.2 Independent Variables ..............................................................................................67
4.7 Questionnaire....................................................................................................... 68
4.7.1 Sampling ...................................................................................................................68
4.7.2 Data Gathering...........................................................................................................69
4.8 Users Profiles ....................................................................................................... 71
4.9 Results and Analysis.............................................................................................. 71
4.9.1 Efficiency ..................................................................................................................72
4.9.2 Task Complexity .......................................................................................................73
4.9.3 Individual User...........................................................................................................74
4.9.4 Effectiveness .............................................................................................................74
4.9.5 Individual User...........................................................................................................76
4.9.6 Users Satisfaction.......................................................................................................77

4.10 Normality Test ........ 78
4. 10. 1 Friedman's ANOVA Test ........ 80
4.10. 2 Wilcoxon Signed-Rank Test ........ 81
4. 11 Discussion ........ 82
4. 12 Summary ........ 85
Chapter 5 ........ 87
Experimental Phase III: The Role of Expressive Body Gesture and Earcons and Auditory Icons in E-assessment Interfaces ........ 87
5.1 Introduction ........ 87
5.2 Aims ........ 87
5.3 Tasks ........ 88
5.4 Hypotheses ........ 88
5.5 Experimental Condition ........ 89
5.5.1 Assistance Types ........ 90
5.5.2 Implementation of Non-speech Auditory Metaphors ........ 91
5.6 Experimental Design ........ 91
5.6.1 Procedure ........ 92
5.6.2 Tasks ........ 93
5.6.3 Independent Variables ........ 93
5.6.4 Dependent Variables ........ 93
5.7 Data Collection and User's Profile ........ 94
5.8 Results and Analysis ........ 96
5.8.1 Achievement Stage ........ 96
5.8.2 Involvement ........ 100
5.8.3 User Evaluation ........ 105
5.8.4 Satisfaction ........ 106
5.9 Discussion ........ 107
5.10 Summary ........ 108
Chapter 6 ........ 110
Conclusions and Empirically Derived Guidelines for Assessment ........ 110
6.1 Introduction ........ 110
6.2 Review of the Experimental Work ........ 110
6.3 Limitations ........ 112

**6.4 Derived Guidelines for the Use of Multimodality ........ 113**
**6.4.1 Description Text ........ 113**
**6.4.2 Images ........ 113**
**6.4.3 Avatar ........ 114**
**6.4.4 Recorded Speech with Images ........ 114**
**6.4.5 Images with Description Text ........ 115**
**6.4.6 Full Body Gestures Avatar with Earcons and Auditory Icons ........ 115**
**6.4.7 Complexity of Assessment ........ 116**
**6.5 Future Work ........ 117**
**6.5.1 Smart Virtual Instruction ........ 117**
**6.5.2 Personalised Virtual Instructor ........ 117**
**6.5.3 Virtual Instructor for Less-able Users ........ 117**
**6.5.4 Other Types of Assistance ........ 118**
**6.5.5 Mobile E-assessment ........ 118**
**6.6 Conclusion ........ 118**
**Appendix ........ 130**
**Appendix A – Experimental Phase 1: The Role of Multimodal E-assessment Interfaces ........ 130**
**Appendix B– Experimental Phase II: The Role of Avatar, Record speech, Description Text ........ 148**
**Appendix C: Experimental Phase III: The Role of Expressive Body Gesture and Earcons ........ 163**

## List of Figures


**Figure 3.1: Snapshots of experimental condition .................................................................. 37**

**Figure 3.2: Users profile in terms of age, gender, education in both groups ...................................... 43**

**Figure 3.3: Total time taken by each user in both groups to answer every question ......................... 45**

**Figure 3.4: Mean values of time by users in both groups and question complexity .......................... 47**

**Figure 3.5: values of percentage correct answers in the VOAP and VMAP ...................................... 49**

**Figure 3.6: Mean total number of correct answers by each user in condition .................................... 51**

**Figure 3.7: Frequency of users' agreement to each satisfaction statement ....................................... 53**

**Figure 4.1: Snapshots of experimental in e-assessment condition ..................................................... 62**

**Figure 4.2: Users' profile in terms of age, gender, educational and prior experience ..................... 70**

**Figure 4.3: Mean values of time taken by users in experimental by each condition ......................... 72**

**Figure 4.4: Mean answering time for questions related to each question ........................................... 73**

**Figure 4.5: Percentage of correct answers achieved by users question complexity .......................... 73**

**Figure 4.6: Number of correct answers achieved by users for all questions ....................................... 76**

**Figure 4.7: SUS Scale results ......................................................................................................... 75**

**Figure 5.1: Profile of users in terms of personal data and previous experience ................................. 95**

**Figure 5.2: Correct and incorrect percentages of answers achieved by users for all questions ........ 99**

**Figure 5.3: The number of correct answers provided by each user .................................................... 99**

**Figure 5.4: Correct recognition of users of the assessment types ...................................................... 101**

**Figure 5.5: Recognition of users for all the tested non-speech sound ................................................ 102**

**Figure 5.6: Results of the user evaluation toward the non-speech sound ......................................... 105**

**Figure 5.7: The percentage of user responses to each statement in the satisfaction questionnaire 106**

## List of Tables

Table 1.1: The outline of thesis and framework of the research ............ 9
Table 3.1: Multimodal used to communicate the information by users ............ 36
Table 3.2: Tack system ............ 39
Table 3.3: Tasks complexity levels ............ 39
Table 4.1: The condition multimodal features ............ 63
Table 4.2: Dependent variables ............ 67
Table 4.3: Independent variables ............ 67
Table 4.3: Test of normality for time ............ 79
Table 4.4: Test of normality for correct answer ............ 80
Table 4.5: Friedman's ANOVA test results for correct answer ............ 80
Table 4.6: Wilcoxon signed-rank test results for correct answers ............ 82
Table 5.1: Mapping among the important level of assessment types and non-speech sound ............ 90
Table 5.2: Auditory icons are used in auditory avatar body gestures interface to indicate users..... 92
Table 5.3: Chi-square values and significance levels relating to the achievement level ............ 98
Table 5.4: Chi-square and significance levels relating to involving test 1 ............ 98
Table 5.5: Chi-square and significance calculations relating to involving test 2 ............ 102

Chapter 1

# Introduction

## 1.1 Motivation

This study will empirically measure the achievement of users, particularly in regards to their achievement; efficiency and user satisfaction, as part of determining the effectiveness of multimodal metaphors (speech and non-speech sound, avatars, and body gestures) within an e-assessment context. It can be stated that one of the motivations of the study is to enhance the user-interface usability, in order to develop the learning experience through e-assessment. Moreover, if technology is to become an integral part of assessment process then there is a real need to identify the potential benefits, in relation to this idea the researcher wanted to examine the impact of the use of the latest technological approaches in e-assessment in relation to the benefits to the user. It is envisaged that the user and those who are responsible for e-assessment development as well as assessors will benefit from an understanding of how multimodality improve e-assessment interfaces to help users.

This research carries on an investigation into the use and effectiveness of multimodal metaphor usage in communicating information and helping users during e-assessment. The investigation's aim is to reveal the appropriate purpose, utilization and combination of such metaphors in order to improve the level of usability in e-assessment interfaces and to enhance student involvement with different types of assessment and content. Online assessment is undoubtedly an important part of the e-learning process, improving due to its ability to deliver convenient feedback to participants, thereby assisting in the improvement of the overall learning and teaching experience. With technological advances in e-learning, it is significant that the corresponding improvements in e-assessment, as a support tool in e-learning, are suited for achieving the overall purposes of using applied science in teaching.

Learning effectiveness can be evaluated through assessment; however, the degree of the assessment effectiveness is one of the subjects which is continuously being posed as a fundamental problem for research. Redecker and Johannessen [1] indicate the fact that there is not enough understanding about assessment and there is a real need to bring technological advances in line with pedagogical approaches. Importantly, the reliability and validity of online assessment techniques need to be reviewed as part of an overall reconceptualisation by teachers of approaches to assessment [2]. Another concern that justifies the need to investigate the effect of e-assessment of learners is that traditional methods may be perfectly suitable and may be preferred by learners [3], thus there is a need to investigate the effect of e-assessment in terms of its suitability and positive effect on the users. Moreover, [3] in recognition of the importance of traditional methods and the emergence of e-assessment, suggests that the combination of these methods may be the solution.

Another justification for the need to investigate the use of multimodal metaphors in e-assessment is that they may have their origins in the traditional methods. Although it is recognised that e-assessment will determine or restrict the types of questions and tasks, it will also offer the opportunity to assess novel constructs [4] and ways in which this can be achieved can be investigated.

It is important that those who are involved in designing e-assessment understand that it needs to be a novel and stimulating [5]. This can only be achieved through understanding the effect of the different design approaches to e-assessment, something that the present study aims to achieve.

Much of the of the literature pertaining to the evaluation of e-assessment does not focus on the aesthetic, but rather on the relationship between advances in e-learning and how they are matched by advances in e-assessment, effectiveness in improving pedagogy or helping those responsible for teaching and evaluation [4].

Therefore, the main challenge related to the evaluation of e-assessment is that there has been a significant technological development for e-assessment with a corresponding concern about its suitability for purpose, even in the areas of graphical design, interface design and the effect of aesthetics. However, in the area of multimodality as a tool in e-assessment there has been a lack of studies, specifically on its suitability and effect in the e-assessment process.

An important issue in the design of e-assessment is that there needs to be consideration of the different levels of ability, in that it should be stimulating to all learners [5]. In light of these challenges, the present study addresses the issue of varying ability by measuring score (performance) and providing different degrees of difficulty in the questioning in the e-assessment exercise, which is measured against the different approaches to using multimodal metaphors.

Both forms of qualitative and quantitative assessment are required to demonstrate and plan learning outcomes; however, it is not by any means a simple task, especially with the condition used in e-learning. Substitution of the human factor in the assessment is not likely; however, some development has been made in assessment systems to perform them electronically.

Assessment in e-learning interfaces is based on the visual and audio channel using either video and text or graphics with symbols to annotate ideas. The intended impact of a single channel communication approach is for the recipients of the information to take further time to understand the information communicated. Users of e-learning and e-assessment systems are classified into two groups; students and teachers.

## 1.2 Aims

- An aim of the study is to investigate the impact and usage of multimodal interaction metaphors of e-assessment interfaces.
- Another aim is to create a set of empirically derived guidelines for the use of multimodal metaphors in e-assessment interfaces to enhance e-learning.

## 1.3 Objectives

- To assess the optimum combination of avatars, body gestures, speeches and non-speech sounds, recorded speeches, images, earcons, auditory icons and describing text in terms of how they affect the usability and users' involvement of an e-assessment method. This will be achieved through a series of experiments.
- To assess the effect on the usability of the different multimodal metaphors, alone and in combination, where usability includes efficiency, effectiveness and user satisfaction. Specifically, efficiency includes the time taken to complete e-assessments, effectiveness include user performance in terms of the overall score and user satisfaction includes enjoyment and ease of use. This will be achieved through experiments where users are engaged in e-assessment exercises as well as a post-experimental questionnaire about their experience using the e-assessment tool.
- To assess the effect of multimodal metaphors on the user for varying degrees of question difficulty. Specifically, to ascertain if multimodality can provide more clarity to difficult questions.
- To provide participants with audio-visual type assessment content during the e-assessment interfaces using creation and use of multimodal metaphors.
- To measure the level of the user's achievements and performance for each condition to show how multimodal metaphors assist students using an e-assessment interface.
- To conduct three experiments developed by the researcher to measure the usability of three different conditions with dissimilar modalities in order to measure the usability of each condition.

- To conduct a controlled experiment using normal text which represents a traditional assessment interface without multimodal. The control experiment also tested for usability aspects, chiefly, user satisfaction, effectiveness and efficiency.
- To general guidelines to improve e-assessment interfaces assisted by the use of designs incorporating multimodal metaphors.

## 1.4 Overall Hypothesis

The overall hypothesis for this study is as follows:

The use of the multimodal metaphors can make e-assessment more efficient, and more effective in terms of the user's performance and satisfaction.

## 1.5 Methodology

The main part of the methodology is to conduct experiments with learners using different multimodal metaphors and then to conduct questionnaires in order to evaluate the effectiveness of different multimodal metaphors and their combinations on the user. As part of the methodology a literature review was carried out in order to reveal how the use of multimodal metaphors can have an effect on users and the various ways this can be measured. The experiments assessed user performance and efficiency and the results from the questionnaires assessed user satisfaction. The questionnaires were distributed to users of varying ages and different backgrounds. All obtained data, both objective (measured) and subjective (user provided) was analysed. A statistical analysis was conducted using a set of tests, examples of which include T-test and ANOVA. The obtained results were dependent upon the specific metaphors tested and the content of the communicated assessment. Conclusions were drawn about the applicability of metaphors and approaches and designs that worked successfully in e-assessment condition on the basis of efficiency, effectiveness and user satisfaction.

**Literature Survey:**

The first step in this research was to review several relevant topics in the literature such as e-assessment, multimodal interaction and multimodal e-assessment systems. This review provided insights into the underlying theoretical background of e-assessment, including e-assessment definitions, pedagogical principles, challenges and limitations of e-assessment. The review also covered multimodal interaction and related experimental findings. Finally, issues related to multimodal e-assessment interfaces and examples of multimodal e-assessment systems were reviewed.

**First Experiment**:

This experiment represented an initial investigation of multimodal e-assessment interfaces by performing an empirical study that was aimed at the evaluation and comparison of efficiency, effectiveness as well as learning performance and user satisfaction of two different e-assessment interfaces. Two independent groups of 15 users (n=30) were involved in performing common tasks and to answering questions related to the presented learning content. These questions were of two groups: The first group of users; control, was provided with a typical e-assessment interface with only text. The second group; experimental, was provided with an interface that combined multimodal metaphors such as describing text, images also facially expressive avatars. Both e-assessment interfaces communicated the same question. The results of this experiment formed the basis to design and conduct the second experimental study in this research. For overall, the first experiment was designed to confirm findings of the literature survey and to carry out an initial evaluation to obtain an overall impression and understanding about the procedure and test criteria.

**Second Experiment:**

This was carried out to investigate the use of avatars as virtual lecturers with images and recorded speeches with images and description texts with images in e-assessment interfaces. The aim of this experiment was to evaluate the usability aspects (efficiency, effectiveness and user satisfaction) and learning performance of three different e-assessment interfaces recruited in the presentation of three conditions. The first condition was avatars as with images and the second condition was recorded speech with images while the third condition was described text with images. Three different e-assessment interfaces were aimed at creating at designs to utilize speaking avatars and images as well as naturally recorded speeches and description texts in order to present audio-visual of the learning materials. The three experimental e-assessment interfaces were tested independently by one group of 30 users assigned to accomplish the required tasks and to answer questions in relation to the communicated learning content.

**Third Experiment:**

On the basis of the derived results in the second experiment, a third experiment was conducted to check the effect of non-speech sound in an auditory message which was used to assist the half-body animated tutor when the assessment content had to be presented. Therefore, as an advanced and extended step of the second experiment, the third experiment made use of the earcons and auditory icons in order to investigate further communication issues of the auditory signals. This condition was called auditory-enhanced virtual tutor with body gestures ABGC. The experimental group consisted of 30 members that evaluated the performance and usability aspects of the ABGC condition in terms of the students' involvement (recall and recognition questions were used).

In the final step of this research, the obtained results from the three experimental studies were discussed as a whole to draw the final conclusions and to derive a set of guidelines to design and implement multimodal interfaces for e-assessment systems. These guidelines are suggested to enhance usability and learning performance in e-assessment interfaces.

## 1.6 Thesis Contribution

The research contributions presented in this thesis have been achieved through the proposition of solutions to the stated problems. The research conducted in this thesis addresses, e-assessment interfaces through the use of multimodal attributes, for example images, earcons, avatars, recorded speech and body gestures as tools to make assessment easier and more effective and to assist in the integration of contents. The results of the study contributed the following:

1. The thesis contributed to understanding the impact that multimodality, in different combinations, has on the effectiveness of e-assessment interfaces. This was in response to a gap in the research where the study addressed the impact of the latest developments in multimodality.

2. The thesis also contributed to producing guidelines that could benefit the designers of future e-assessment interfaces. Specifically, ways to integrate earcons and auditory icons in the e-assessment interfaces so as to encourage the role of the body gestures are contributed as well as scope to judge their performance within interactive e-assessment in the context of modern learning.

## 1.7 Outline

There are six chapters followed by three appendices in this thesis. A short description of these is given below.

| |
|---|
| **Chapter 1**<br><br>Introduction, Aims and Objectives, Methodology, Motivation, Overall Hypothesis & the Main Contribution |
| **Chapter 2**<br><br>Literature Review: E-assessment, usability and Multimodal Interaction |
| **Chapter 3**<br><br>Experiment I<br><br>Control group (15) users **(VOAP)** experiment group (15) users **(VMAP)** |
| **Chapter 4**<br><br>Experiment II<br><br>Avatar with Images (AI) condition (3 Tasks) \| Record speech with images (RI) condition (3 Tasks) \| Description Text with images (DI) condition (3 Tasks) |
| **Chapter 5**<br><br>Experiment III<br><br>Auditory and Avatar body gesture condition<br>One group 30 users 9 Tasks |
| **Chapter 6**<br><br>Conclusions<br>A set of guidelines |

**Table 1.1: The outline of thesis and framework of the research**

# Chapter 2

# Multimodality and E-assessment Interface

## 2.1 Introduction

The Literature Review is one of the research methods employed in this thesis. This section will review and analyse the work of previous researchers relevant to the research conducted in this thesis. This literature review covers many aspects, with emphasis on three major sections; defining the key terms of e-assessment; multimodal interaction; and multimodal e-assessment. Moreover, this allows for the provision of a theoretical basis to assist the research as well as facilitating the determination of its nature. In this study, theoretical, experimental and practical analysis will be carried out based on the interactions present within a multimodal condition of e-assessment. This chapter presents an overview of several theories of a significant nature which have been explored in depth by the researcher, with a view to emphasise the core objectives presented by this thesis. This has been categorised and arranged into three main sections:

The first section provides preliminary information about e-assessment definitions, benefits and challenges as well as the limitations of e-assessment.

The second section provides the fundamental concepts of multimodal metaphors for use in this research, specifically visual metaphors, speech and non-speech sounds, and also avatars, in order to provide information about multimodal metaphors to improve users' computer interaction and identify a variety of solutions to resolve some problems.

The last part of the review focuses on a critical review of multimodal e-assessment interfaces with a close look at the research studies conducted in this area about enhancing usability.

### 2.2 E-assessment

Improvements and innovations in the area of information technology have been able to offer new assessment implementations and methods. Learning and assessment are undoubtedly complementary to each other. Information technology has a deep effect on e-learning and e-assessment systems. Computer-mediated assessment, computer-aided assessment, online assessment and e-assessment are some of the connected terms used in relation to information technology with assessment [6]. Assessments will be conducted for two purposes: to assess students' progress and to assist web based student learning. Assessments may also be used for course and tutor evaluation. Noting and reviewing teaching strategies and web-based material may lead to enhancements and afford alternative means for effectiveness [7]. The design of the assessment should be one of the first considerations when creating an online course. The assessment should be an integral part of the program and should not be considered as a separate aspect [8]. Enhancing the quality of the student learning experience is a major issue in the education sector, and it has been extensively recognised that e-assessment can contribute to this. Nevertheless, it is interesting that while much research has been carried out into the attitudes of e-assessment by a fraction of instructors, e-learning experts (Bull&McKenna, 2004; Stephens&Mascia,1995; Warburton & Conole, 2003) [9], believe that there is comparatively little research into what students believe. While we often make assumptions towards what students feel, it would be helpful and interesting to put these to the test and get some online data from students themselves [10].

Both Formative and Summative assessments make up the types of assessment methods used to measure the users of e-learning. By "ranking students to show learners' accomplishments", the use of the summative assessment method allows for the provision of a final ruling of the achievement made by students in relation to the specific objectives, while the other method known as formative assessment is used as an "analytic tool for students and teachers to determine and enhance areas of weakness" [11].

Clark [12] recommends seven principles when considering pedagogical aspects. These principles provide a mixture of both emotional and cognitive educational components which will present a superior result upon incorporation and facilitate e-learning in becoming more efficient, and making an enhanced experience for the students. To be brief, these are very important skills which are required in order to keep things light, examples of which include establishing the involvement of an emotional nature, linked concepts, pragmatic practices, elaborated examples and also repeater reflections.

### 2.2.1 Definition of E-assessment

An e-assessment method has been developed which is based on a multi-dimensional approach, which includes student friendly services and a user central nature that accurately follows the regulation of Human-Computer Interaction (HCI) instead of Human-Human Interaction (HHI) [13]. According to the JISC e-assessment is defined as follows: *'The end-to-end electronic assessment processes where ICT is used for the presentation of assessment activity and the recording of responses. This includes the end-to-end assessment process from the perspective of learners, tutors, learning establishments, awarding bodies and regulators, and the general public'* [14] . Jordan [15] provides a wider definition and says that it includes the use of a computer for assessment activities, whether they be summative, formative or diagnostic. A more simple definition is provided by Stowell and Lamshed [16] who argues that 'E-assessment is the use of information technology for any assessment-related activity'.

### 2.2.2 Benefits

Using information technology in e-learning strategies provide an effective way of assessing both teaching and learning through supporting traditional ways of assessment, and one of the associated advantages is that it yields rich data that will help educators to further understand both teaching and learning [17]. This idea is supported [18] who says that e-assessment provides convenient feedback to all participants which helps to improve overall teaching and learning. The development of e-assessment technology has in fact opened up new possibilities in terms of new constructs that can be used to assess learners Boyle and

Hutchinson [4]. These authors argue that e-assessment can be used for more sophisticated assessment tasks, in that the type of questions could be different from those used in paper assessments; examples of such include case studies and simulations. Boyle and Hutchinson [4] refer to this as a sophisticated e-assessment and say that it has benefits over previously used methods. The literature speaks about e-assessment in reference to the learner and the instructor. Sainsbury and Benton [19] indicate that both teachers and pupils respond positively to e-assessment, for the teachers, they are able to access both qualitative and quantitative reports which help teachers to organize their teaching according to the results. According to Ball [20] e-assessment provides the assessor with more tools and improves the accessibility for those being assessed, particularly in relation to improved accessibility and usability for those who are disabled.

Certainly there is the idea that e-assessment improves the quality of learning for students, and according to a study by Dermo [21] students themselves agreed with the benefits that e-assessment can bring to teaching and learning. However, Dermo's [21] approach is different because he examined the attitudes of the learners rather than the instructors. The idea that the user or more specifically, the student, acknowledges the benefits of e-assessment is acknowledged by Hodgson and Pang [22] who found that there was high uptake and high satisfaction of formative e-assessment among students, and according to these students it even lead to a change in the learning approach, specifically encouraging students to find the answers.

### 2.2.3 Limitations

One of the main limitations that could be associated with e-assessment is:

- That it provides a greater opportunity for those being assessed to cheat. Prakash and Saini studied the use of an online quiz by students and included academic honesty as one of the variables, and they found that over half of the students received help from other students during the assessment. This also brings attention to the fact that students may also refer to other sources for help.

- Prakash and Saini [23] also agree with this limitation and emphasize the role of plagiarism detection tools used in e-assessment [24], and Mothukuri et al. [25] suggest an approach that involves one place and one set time assessment in a ubiquitous environment in order to reduce the likelihood of malpractice.

- Wielicki [24] suggests that students involved in online classes, as opposed to purely online assessment, are less likely to cheat because they are too overworked and unorganized; however, there is a need for skillful instructors to achieve this.

- E-assessment is that often it does not accommodate different types of students in terms of their ability; Prakash and Saini [23] argue that some students may have writing skills and other students may be more creative or good at problem solving.

- The technology itself may be a limitation of e-assessment, in a study by Caniou and Excoffier [26] it was suggested that in order for teachers to provide a new type of questionnaire would require them to have computer science knowledge. This need for teaching staff to have the required technical expertise due to an increase in the use of technology in teaching and learning is echoed by Dika et al [27] who indicate that institutions of learning now have to deal with new issues.

### 2.2.4 Challenges

Looking to the future of the use of technology in assessment, according to Redecker and Johannessen [1] we are now at the junction of two different assessment paradigms and there insufficient pedagogical commitment and understanding to move from the era of computer-based assessment to an era of embedded assessment. However, Redecker and Johannessen [1] also recognize that it is not only about pedagogical commitment, it is also about the technology because they argue that technological advances have to be in line with new assessment paradigms. Similarly, Gikandi et al [2] say that teachers should reconceptualise issues of teaching, learning and assessment in light of non-traditional assessment approaches offered by technology, ways of doing this include reviewing the reliability and validity of online assessment techniques.

In relation to this idea teacher commitment, McCann [28] brings attention to the fact that in order for an e-assessment system to be adopted by faculty members it has to fit with faculty culture, unfortunately, in their study it was found that a centralized e-assessment system was not accepted by faculty members because they preferred to work autonomously and a centralized e-assessment system requires cooperation. Therefore, one of the challenges facing the adoption of e-assessment in higher education is the lack of change strategies in institutions.

According to Llamas-Nistal et al [3] there should not be the assumption that e-assessment is the panacea for all assessment situations because students may prefer to provide answers using a traditional pen and paper rather than using a keyboard. Llamas-Nistal et al [3] do offer a solution or compromise where they say that there should be a combination of e-assessment with traditional assessment techniques. Moreover, one of the key challenges in e-assessment is to not allow the developed technology to dictate assessment practices. In relation to this idea Boyle and Hutchinson [4] indicate that as e-assessment develops it will have a significant effect on questions and tasks; however, as mentioned above they can also provide a way of assessing novel constructs.

An interesting challenge for the future of e-assessment has been put forward by Johnson-Glenberg [5] who says that instructional designers have to design e-assessment so that it is stimulating and novel to a wide range of learners, that it should acknowledge that users may have different levels of knowledge and therefore, they should be given more control through optional navigational links and paths. In relation to this idea, Stodberg [29] argues that it is important that e-assessment tests are fair and should not present a disadvantage to some students because of the e-assessment procedure. This issue is also raised by Al Smadi et al. [30] who emphasizes that students have different learning styles which present a challenge for educators because they have to develop e-assessments suitable for appropriate all students. Overall it has been put forward by Hodgson and Pang [22] who, although agree that e-assessment encourages learning, especially in a learning community, there is still further to go on nurturing this type of learning.

In reference to the evaluation of e-assessment, there are many areas that have been addressed. Because e-assessment is seen as a development from traditional methods, it is evaluated against such traditional methods and are criticised against the merits of traditional methods [31]. The idea is that the use of technology, if used with imagination and skill, can add value to assess and that e-assessment is not just another way of doing assessment but can in fact increase learner performance [32].

One of the advantages of e-assessment is that it provides immediate feedback to learners and this has been evaluated in relation to improving functionality and usability [33]. Often, e-assessment is evaluated against the learning process itself or how it forms a part of the learning process [34] or how it can facilitate learners with different abilities and assist those involved in pedagogy [35]. Moreover, there has been an evaluation of e-assessment approaches in terms of their need to keep up with advances in e-learning and how the two can be developed together using technology [23].

### 2.2.5 Usability Evaluation in E-assessment Interface

Usability is defined as the effectiveness, efficiency and satisfaction with which a user can attain a particular aim in a specific environment. Measuring and assessing usability have been considered as an important part of the product or method enhancement process. It can happen at any level of the improvement and execution, and lead as either a formative or summative exercise [36] [37]. It is believed that the insertion of usability characteristics within the e-learning instructional design method is vital to an effective e-learning curriculum. This inevitable as can be seen from the works of [38], [36] and [39] which provide that there are six general attributes of effectiveness in usability including utility, learnability, efficiency, retain ability, errors and satisfaction. The insertion of usability as part of an assessment process has developed the quality and effectiveness of web-based directives [36]. Multimodal learning is known as the incorporation of text, audio and visual elements in one [40] condition. A multimedia learning method interacts with the users to communicate information that assists the learning system. Hansen [41] refers to these methods as interactive communication, the purpose being a focus on the creation of forms of digital media, including text, line drawings and animation, as well as audio and video.

The area of evaluating the user interface in e-assessment is not only approached from the viewpoint of the effect of multimodal metaphors but also from the point of simply aesthetics. Enhanced aesthetics have been shown to improve the factors that are being addressed in the present study, namely, participant satisfaction, time taken to complete tasks and task performance [42]. However, the method that was used by Miller for the experimentation on the effect of aesthetics on these factors focussed on the aesthetic and seemed to neglect the context which was the purpose of the e-assessment itself. While it is important to address how aesthetic enhancement has an effect on the user performance, any experimentation should take place using a real e-assessment format, since the overall aim is to address aesthetics and its impact on the learner in e-assessment. Miller’s study focused on the aesthetic and the experiments using photo, description, story retelling and picture naming may not have been realistic or representative of e-assessment where a number of different skills are being tested, not just cognitive capability against the enhancement of aesthetics.

The objective of a usability valuation is to evaluate the quality of a learner interface design and create a basis for enhancing it. Usability assessments and related activities assist designers in making improved decisions, and help them do their jobs more effectively [43] and [44].

The product which has importance is the method itself. A usability valuation which focuses on the final method is usually denoted as a validation exam. The most typical product that is used for a usability valuation is still incomplete, and only exists as an operational model of the technique. This is usually denoted as a valuation test [45]. Other forms contain meetings with designers, edited videos as well as re-design suggestions.

Cooperative design operations are the best methods to involve all representative stakeholders, particularly representative learners, in the design and improvement operations early on in the enhancement series. When used in the early stages of the enhancement series, feedback collected from these techniques will contribute to the improvement of a user centred website with the least quantity of potential and cost.

The possible risk is that representative users can become so recognisable with the condition design that they start to envisage further just as design and improvement specialists and less like target learners [46]. Cooperative design requires additional time for stakeholders. Carrying all stakeholders in a meeting could impact the input from users who might be reluctant to offer ideas when designs are presented.

## 2.3 Multimodality

In overall terms, the literal sense of mode is the technique through which a certain work is accomplished. The term "multimodal" refers to accomplishing a task via the use of a number of methods all combined together. Multimodal in its most fundamental meaning is the coexistence of more than one semiotic mode within a specified context. Furthermore, multimodality is a daily reality. It is the knowledge of living: we experience everyday life in multimodal terms, by sight, sound, movement and gesture. Consequently, in today's world, consideration of multimodality in academic investigation seems to have become even more important and even more related [47]. For example, information that is used for the electronic assessment methods can be seen in the form of description texts, audios, videos and also images. Thus, all these are diverse ways of representing the same or dissimilar information. Collecting all these ways into a single communication in a synchronous or asynchronous way result in what is called a "multimodal" form of representing the information for the assessment processes. In other words, the majority of the computer scheme developers concentrate on the visual modal only for information transfer; thus this may be confusing to the users and affect their attention because of the overloading on this channel through the interaction [48]. In this study, the term multimodal metaphor is used to indicate the use of auditory and visual metaphors to represent the information to be used in the online assessment methods.

What is the most important thing I must do if I want to make the interface easy to use? It is not "Nothing Important would ever be more than two clicks," or "Say the user language," or "Be consistent," [49]. It is important for the user to obtain what they want from the computer during the interface. As we know, users [50] deal and interact through their sense, which varies depending on every person. In computer sciences, the sense of the term “modality” is ambiguous. In human computer interaction, the term typically refers to the human senses: hearing, touch, smell and taste, but several researchers distinguish between computing modalities and the sensory modalities of psychology. Sharon Oviatt presented a more practical description, demonstrating that multimodalities for example speech, touch, hand gestures, eye gaze and head and body movement are multimedia schemes of output” [51] and [52]. Nigay and Coutus argue that “multimodality is the communication used by users in a variety of sense channels to be delivered systematically” [53].

A multimodal interface is a human-machine interface that collects multiple channels of communication between user and machine [54]. The model most frequently used are a combination of gestures and speech (accompanied by a gesture of designation) and interaction with both hands on an interactive tablet. Thus, we can conclude that multimodality focuses on using tools to communicate among the user and interface, indeed more so than using it as a one method channel as it is used in the traditional or typical sense. As stated above, the user can only use one of the senses to interact with the computer monitor. By only using one of these senses, this will mean a lack of use in the rest. Moreover, there are other causes listed below that let us produce multimodal:

1. Data overload [55] means that the user becomes confused when they are offered a huge quantity of information which is often the case when only one channel of communication is used. The user typically uses one sense [56] but fails to use the other senses because of data overload.

2. Enhancing performance of recognition-focussed schemes [55] means the interface may be used in a flexible manner that produces an intelligence scheme. Furthermore, it improves the performance of recognition which reduces errors [57] effectively.

3. Greater feeling of immersion in virtual-indeed environment [55]

4. Support time involvement and attention management of the difficult real-world [55]

Furthermore, it is essential to make computer technology further usable by people. The designer of an interaction scheme should have knowledge in psychology and cognitive science [58]. There are two sets of multimodal interfaces, the first of which is where a group allows several user insertion modes together, for example speech and pen. This allows the user to improve his experience, power and ability. The second group lets the user [55] use a multimodal scheme insertion with visual and auditory responses.

There are some guidelines for the improvement of multimodal schemes which have been built on the analysis of the particular quality that sensory channels have. The guidelines which represent the major decisions and consideration engaged in the procedure of designing a multimodal interface are [55]:

1. The choice of modalities: some causes have been offered for including several modalities in the design of an interface. Many researchers recommend that it must be supported to account for variances in user favourites, needs and abilities. Furthermore, the choice needs to consider the tasks and kind of information that the user can handle.

2. The creation of modalities in relation to tasks and types of information: the creation of a natural design among modalities and the information and tasks to be offered.

3. The combination, synchronization and integration of modalities: this is a sample that clarifies that one often-employed modality order is the use of an auditory alert, followed via the visual presentation of related information. “Cars use an auditory signal to notify the driver of a forthcoming turn, and a visual display then offers more detailed information”.

4. The adaptation of multimodal information display: “multimodal interfaces need to be flexible and take into consideration possible changes on the needs and abilities of the users, their tasks, workload and environment that they are operating in” [55].

The use of multimodality is assumed to be more beneficial than the use of a modal alone. These benefits support the users whose aims are to use the interface. The following are some benefits:

1. They diminished errors which might happen via users.

2. Allow the interface to become clearer.

3. It is easier to determine what has been going incorrect

4. Allow for more bandwidth to the communication.

5. Increment alternative communication systems to diverse situations and environment.

With respect to multimodal interaction: a multimodal HCI scheme is only one that responds to the insertion of more than one modality or communication channel (record, gesture, writing and others) [58]. With regards to the rating of vision techniques for MMHCI per human body - wide body movements, there are three significant issues in articulated motion analysis: (a) Representation (b) Computational paradigms (c) Computation reduction. On the one hand, gesture recognition, as per psychology studies, indicates that there is a relation between conceptualising capacities and our language abilities. Gaze is defined as the direction to which the eyes are a strong indicator of attention [58].

## 2.4 Visual Display

A visual presentation is a channel which is used to display objects generated in the computer by an interface. The vision sense is used to display an object in a space and can also be represented by text or diagrams [59]. In visual modality, the designer is considered to be using a common graphic for example more images or graphics [60]. Examples include the Rational Rose [61] software which makes a visual model by converting a user's class diagram to class directly in UML linguistic. Essentially, there is a good impact achieved by using a visual model because it allows for a smoother and easier interaction between the user and the interface. On the other hand, using more images and graphics in the interface might cause overloading and confuse the user. Therefore, it is imperative that the interface be made simpler in order to meet the objective of refining user interaction. With visual presentations, the user should be keeping his/her eyes on the interface so as to focus on what is going on. Nevertheless, auditory channels will make a contribution to other senses.

Visual show is accountable for representing the required information in a manner so that it can be seen visually and it is achieved by a range of methods for example, via designing a web page (the world wide web) and representing the information on it, creation of graphs, tables, images, or offering a summary in pictures or text; all are the methods via which the useful information can be shown to the required individuals. In actual fact, using the visual displays is the easiest and most efficient methods of representing the information [62]. Nonetheless, such strategies are useless for individuals possessing visual disabilities and require specifically designed methods and ways for information communication and understanding.

### 2.4.1 Auditory Modal

This paradigm deals with the representation of information in a system so that it can be easily explained and perceived via the human auditory senses. The auditory exemplar is typically represented via the speech.

### 2.4.2 Speech

A speech modal is a channel that is used to represent particular information to users using voice [63]. Natural speech can be classified as usual speech spoken by humans, which has been recorded via particular software and stored as digital files. When played, it offers natural human-like interaction with computer schemes [52]. The use of sound in an auditory modal as an alarm has been proven. In fact, this is primarily focussed on using an indication to bring attention to the user that there is something wrong or right. This was first practised or used via [64] and [65]. As was mentioned in the former section,  the overloading of information might be beneficial for schemes having a limited screen area. Therefore, the presentation of some information in sound will assist in decreasing the amount of text and graphic required in the interface [63].

Also, this will utilise other senses, such as hearing and sight. Mountford and Gaver [66] suggest that voice is useful because the sound is well-known and common for processing information which is used in the user's daily life. There are several ways of utilising sound, such as supporting an object which is focussed on a visual model where the sound can be used for instructions, thereby directing the user [67]. Examples of this include the car navigator, where the sound is used to inform the driver about what is drawn on the map. It is important to have some information regarding the perception of sound [63]. This explains the relationship between the features of voice that takes place inside the ears and the feelings which are generated in the auditory schemes.

### 2.4.3 Recorded Speech

The speech comes first between the metaphors that can be used simply without much hardship and effort because of the human computer interaction. This is the case because a simple microphone followed by an analog to digital converter is used to make the interaction of the speech with the computers. More recently, computers have been using speech recognition software and fixed algorithms that insert the speech signal of the human and convert it to the digital form understood via the computer [68, 69].

### 2.4.4 Non-speech

Non-speech sound metaphors in auditory channels are non-verbal cues that transmit information around objects in the computer interface. These can be made of digitally recorded or synthesised musical instruments, everyday sound effects, or electronically produced pure tones [70], [71], [72]. Published in 1989, the particular issue around non-speech audio of the Human–Computer Interaction journal set an example for the auditory channel theory and practice in computer interfaces [73]. Further techniques of non-speech sound applications in computer interfaces related to the use of earcons (abstract musical tones that convey information around procedure, events, or objects at a computer interface). It is significant to observe that the mappings among the information and the earcons should be learned, because the sounds do not have a direct meaning related to the represented information [71]. The methods of addressing and delivery of the information around a product cause the user to be bombarded with a surplus of information known as information overload, and this situation is difficult, especially when dealing with the introduction of new strategies and systems such as making use of multimodal interactions and metaphors [74]. To illustrate this, it can be observed that an MCKMS mode works to deliver the information and knowledge around the product by collecting speech, environmental sound, and metaphors focussing on the rising pitch. In another instance, the ACKMS scheme combines the speech, earcons, and avatars to deliver the knowledge and information [75]. Therefore, the use of multimodal metaphors allows the users to save time and acquire the requested information in an efficient and fast method when compared to the technique of communicating information using only text and graphic in a combined method. Although the use of audio-visual interfaces display a better choice to the users searching for the information, it has had severe problems when the communication of the audio messages took place as previously shown through the experimentation phase. Later on when the users had knowledge of the scheme, the performance of the scheme guaranteed the user's satisfaction. It was also found that the diverse components of the multimodal metaphors had diverse performance when attempting to communicate the knowledge. An example of this is where the performance of the auditory icons was greater compared to other components of the multimodal metaphors whereas communication of knowledge by sounds was found to be familiar with the exterior sounds in the environment. Earcons, however, were found not to be used as much as the auditory icons in communicating knowledge, yet they showed to be better for communicating knowledge. Moreover, experiments proved that recording and summarisation of the speech, metaphors and a collection of them during the communication of the respective long and small messages was a promising idea with beneficial outcomes. Moreover, models for combining the recorded speech with other modalities have shown development in the performance of the CKM scheme with improved and raised interactions between users of the scheme [76]. In other words, it is valid to say that it is a good idea to use the multimodal metaphors for delivery of knowledge and information around the products and bring development in the E-CKMS schemes and its use, which can be confirmed from the research achievement in the Specialty as software engineering, Internet browsing, and e-commerce [76].

### 2.4.5 Auditory Icons

There is a growing demand for research that recommends merging non-speech sounds (earcons and auditory icons) with graphical interfaces to decrease the visual workload which impact the users' performance [77]. According to [78] auditory icons are defined as "everyday sounds mapped to computer events by analogy with the everyday sound producing events". They provide a method that sounds natural in representing data that is dimensional and also represents the objects that are conceptually in specific computer schemes. The auditory icons allow the data to be categorised into different sets using a single sound [79]. One of the most important advantages of using these is that the sounds used in them are those which people hear in their daily lives, and link them with a specific action [80]. An example of this can be found in the virtual world where we would hear the sound of an object crashing into a wastebasket when the file is deleted, or marked for deletion. This category of auditory icons is like the sound effects which complement the visual events with an appropriate sound in a computer scheme. Yet, their purpose is not just simply to serve as entertainment tools, but also to convey very important information regarding the events taking place in a computer scheme – this allows the user to listen to the sounds from a computer as he does from the everyday world.

Systems like EAR (Environmental Audio Reminders) play a variety of the non-speech audio cues for offices and the common areas within EuroPARC in order to keep us up to date regarding the various events taking place around its building; Share Mon utilises background sounds in order to spread awareness; Sound Shark, the sonic finder, is useful when incorporating the auditory icons in an interface that is well known and used often – the simplicity of it leads people to underestimate the functions that auditory icons are capable of. For this reason, Gaver and Smith [81] demonstrated auditory icons used in a large-scale, multiprocessing, collaborative system called SharedARK, and called the resulting auditory interface SoundShark [82]. Whoever [83] said the analysis of both source and sound are not usually significant, although that [83] has introduced an ad-hoc synthesis to let users recognise sound instead of the analysis of source and sound. These systems display the extensive range of functions performed by the auditory icons. These include provision of information regarding the user's actions, the possibility of new actions and also the object's attributes that are not visible in the system. They also provide the background information regarding the modes as well as processes in a system that is more complex.

### 2.4.6 Earcons

Tuuri et al [84] define earcons as, "nonverbal audio messages used in the user-computer interface to provide information to the user about some computer object, operation, or interaction." Examples of computer objects in this context include menus, files, and prompts. Examples of operations include compiling, editing, and executing. An interaction between an object and an operation can be exemplified by the editing of a file. The earcons are short, non-speech, musical sounds that are used in the interaction processes between computers and humans, and their job is to convey and communicate the information about different objects, operations, and the interfaces involved within the human-computer interaction. “An earcon is defined as a combination of musical notes, called motives, or even a single one, with specific characteristics, such as changes in duration, tone/timbre and loudness” [85]. Earcons are associated with either objects or actions presented in a computer interface. Because earcons make abstract associations with data, users must learn them in an initial training process [86]. Earcons related with device palettes were examined in a computer drawing program to emphasise its usability [87]. They are constructed from the short term musical tones and can be further made shorter and in this way they can be used to convey and communicate the information about the complex systems. A number of experiments have been accomplished in different domains to check the information conveying contents of the earcons. All the experiments have proven that earcons are the best form of communicating information within sound signals.

### 2.5 Avatars

An avatar is a new function that is used in the interface to interact with a user and represents a real human being’s face as a graphical image of a user [88]. The avatar can be either the head of a man or woman, or a whole body. So it is an image which represents the expression. The idea behind the avatar is to simulate a user by using an actual human. When the user is in front of the screen of a computer, it may not be very interesting if the communication is largely human to computer. To combat this, the avatar combines all modal senses, “visual and auditory”, in 3D. The link between the user and the data is the avatar [89]. The avatar is used in many fields. As we can see, it is used in computer games, ATM machines, advertisements and e-learning. It is noticed that people spend a lot of time playing games because they find there is something interacting with them. Also, in e-learning, it can be used as a lecturer to teach the students and to represent information on human activity [90, 91]. Furthermore, as observed in business, it can be used to give information about a new product with brief explanations. The avatar is not a video clip; it is built on interactive elements. That means the avatar communicates with the user, reacting to the user's requests [89] in a clear manner.

Avatars were categorised as naturalistic, abstract or realistic in form according to McLaughlin et al. A feature of abstract, otherwise known as symbolic avatars is the ability to represent the real users whilst maintaining complete anonymity. It is therefore not recommended, due to the fact that it falls deficient in providing a user-friendly environment to enrich the experience of the user which is expected from communication through multimodal [92]. Recently, the avatar became commonly used in many aspects to develop interactivity – both learning engagement and cultural factors are important design considerations [89].

It is another kind of the non-speech multimodal metaphors that combines the use of audio and visual senses in the interactions of humans with computers. Since it combines the two senses, all the advantages of the audio and visual metaphors are combined in these metaphors. In general, avatars can be classified as abstract, realistic and naturalistic. Abstract avatars are cartoon-like interactive characters with limited animation. The help avatar embodied in Microsoft's 'Office' applications is an apparent example of these avatars, designed to provide the users with helpful information during the preparation of their documents [90, 93]. Realistic avatars offer a real representation of humans being generated based on captured static or video images and are used in several applications such as games, movies and teleconferences.

### 2.5.1 Body Gestures

Non-verbal messages, communicate a significant amount of information [44, 94]. Although body gestures are culturally dependent, strong messages of emotion and attitudes are communicated [95]. Body gestures in avatars are used to enhance speech and add emphasis [44, 96]. By using our hands, heads and feet, we can represent a very wide range of signs, signals and movements [95] For example, instead of calling someone, we can use a hand to point to him/her; also nodding the head means it is agreed to something and so on [96]. However, occasionally some gestures can become confusing, so the culture and the context define the meaning of gesture [95], such as a person who taps his temple with a forefinger may be implying that the other person is intelligent or crazy. Basically, using body communication is highly recommended because it sends a strong message that it can emphasise personal feeling or a specific object.

### 2.6 Multimodal Systems Interaction

The assumption of a collection of information pertaining to multimodal features in human beings exists as established by human cognitive studies, and this can take place through the conversion of all raw input data, as extracted via auditory, visual, and tactile sensations. The process thereafter coordinates the flow of information, facilitating the production of cognitive responses which can be perceived. The natural method as used by Hamas in unconscious communication exercises is multimodiality which makes use of a diverse range of information through various channels. It is possible through multimodiality for humans to move, listen and speak at one time [97, 98]. However, most of the computer scheme developers focus primarily and sometimes exclusively, on the visual sense to carry information; consequently this might cause confusion to the users and limit their retention as a result of overloading the visual channel through the developmental interaction [48].

### 2.7 Why Multimodality

Improvements to the interaction between machines and humans can be observed through the linkage of multimodal metaphors to more than one channel, allowing for the conveyance of diverse information [99] which subsequently reduces the memory capacity employed [100]. Dependent upon the speech, setting, touch, gesture or movement, the use of input and output styles will be allowable due to the method of adapting the elasticity environment to interact with computers [97]. It is a characteristic of a successful multimodal interaction that problems faced by the users during interaction with e-learning schemes are overcome, and users are able to base their understanding on the content rather than the technology itself. Instead, users may benefit through the cultivation of their knowledge of technology interaction combined with this as a basis for focussing on learning [48]. As a further incentive, computer-human and human-environment interaction can be made more natural when assisted by human-human interaction, [50] which may help overcome the issue of the absence of face-to-face communication as found in computer user interfaces [96, 101]. It is also observed that without contribution via a common visual input or a signal gesture, spoken utterances are typically unclear [97]. Consequently, human-to-human messages are enriched through the inclusion of nonverbal activities, examples of which are gestures and facial expressions. Furthermore, information similar to what is intended can be offered through diverse channels via multimodal interaction, [99] thereby allowing users alternative ways for computer application interaction in line with that which is more compatible and appropriate to their abilities, requests and preferences [102]. After careful examination of a diverse combination of earcons, avatars and speech, these multimodal were found to have delivered effective outcomes and assisted the development of the interface usability as well as the performance of users [96, 103, 104]. Finally, another use of multimodal metaphors was identified through assisting disabled people in their communication with the interface, facilitating the process and displaying a more improved outcome than before [50].

## 2.8 Humanising Interfaces

Humanising interfaces have long been one of the fundamental purposes of majority of thesis' pertaining to Human Computer Interaction. Humanisation has two objectives; to make the interfaces easier and more enjoyable to use and to make the interface more similar to humans [105]. The process of anthropomorphism offers interfaces to computer schemes via the provision of some human-like characteristics. This is predominantly achieved using speech output, the ability to identify speech, by speech recognition, providing examples of speech and kinaesthetic feedback, social cleverness and the possibility of identifying faces [106]. On another note, researchers are focusing on developing an understanding of the technological schemes essential to integrate some or all of these diverse methods, particularly understanding the output of the computer with specific reference to the designs of interaction that are used in these human-like interfaces [107].

To illustrate the basic concepts and some of the problems mentioned above, we'll use the overall research of the Human Computer Interaction, according to which, there are two participants in the interaction, man and machine, which are seen as two separate agents. They are physically separated, but are able to interchange information by a series of channels of information. As we saw in the first chapter, there are two methods that include the human user: the perception and interaction [108]. Within the methods of perception and interaction that take place in a multimodal interface, you can determine diverse stages of epistemological observation, related to both the man and the machine.

### 2.9 Multimodal in E-assessment Critical Review

Turk [109] argues that our interaction with the world around us is multimodal because we use multiple senses in parallel or sequentially and we perceive the world through stimuli using these senses. However, this is the human interaction with the natural world and Turk [109] indicates that human–computer interaction has traditionally been unimodal whereby people gain information through a single channel. However, this idea that the human interaction with the natural world is unimodal is not reflected in the ideas put forward by other writers. Alharbi [110] says that there are three modes, namely; visual, audio and tactile. The visual is the use of sight, the audio the sense of hearing and the tactile is the sense of touch. When more than one of these modes is used in interaction, then it is referred to as multimodality and it can be argued that normal everyday face-to-face social interaction is multimodal because it involves the cooperation of visual-spatial and vocal modalities [111].

Dumas et al. [112] say that multimodal interaction has allowed our interaction with computers to be more human because it involves the use of speech and gestures and that multimodal approaches have produced a more reliable interaction between human and computer. Dumas et al. [112] also say that people prefer multimodal over unimodal interaction and it improves handling and reliability; however, efficiency is not an advantage of multimodal interaction because there are only marginal increases in this area. Lee and Spence [113] also showed that there is an increase in task performance and longer term benefits to using multimodal feedback over unimodal feedback.

## 2.10 Summary

This chapter describes the significance of multimodality, e-assessment and the need for multimodality. Multimodality means, including a number of senses such as the visual, audio, and video (visual plus audio) into a single and unified mode of communication for clearer user interaction. The speech can be natural or synthesised. The use of natural speech is clearer when compared to synthesised speech. The earcons are non-speech short musical sounds that are used to communicate information around diverse objects, and other events in the interfaces they can be used to communicate better volumes of information in a user interface. E-assessment should be conducted in such a method that it should enhance the learning capabilities of the students as well as ease the process of human-computer interaction. The primary aim is the enhancement of the usability of the user-interface as apparent in e-assessment systems, through assisting users and improving levels of learner concentration, allowing users to experience a renewed learning method in an endeavour to explore the influence of multimodal metaphors on assisting in and improving the overall levels of concentration as experienced by learners.

This research includes the implementation of three primary experimental interfaces (conditions), each of which demonstrates integration with a selection of multimodal features, diverse in nature, examples of which included: speech, description text, images, earcons and avatars. The objective was the measurement of efficiency, user satisfaction and effectiveness. The following three chapters outline the outcomes acquired from the experimental studies, comparing and discussing them in order to create conclusions of an empirical derivation with a focus on identifying the most preferable interface as a yardstick for e-assessment applications from the interfaces which have been examined.

Recently, e-assessment systems using multimodal metaphors have gained significant importance. It was clearly revealed that using one kind of channel to present assessments on the e-assessment interface decrease the usability of the interface. The experimental conditions as described in the following chapters will demonstrate effective solutions to encourage user involvement in the e-assessment interface and also reduce the time spent to achieve satisfaction with the assessment.

# Chapter 3

# Experimental Phase I: The Role of Multimodal E-assessment Interfaces

## 3.1 Introduction

Due to the lack of research available on using multimodal metaphors to increase usability of the assessment interface, it is important to investigate the use of multimodality to engage users with the assessment.

This chapter explains and empirically describes the experiment used to examine the usability aspects of e-assessment interfaces within an e-learning framework that, incorporate a group of model texts with multimodal (avatars, images, description text). The idea is to verify what is the most suitable multimodal metaphor to be used in the e-assessment interface to achieve the highest level of usability?

Furthermore, this chapter addresses and explains two interfaces whilst utilising the same multimodal metaphors that other research groups have used. The first experiments are testing in e-assessment without multimodal (control group); and the second interface investigates the role of multimodal (description text, images, and avatar) as used in the delivering of information in e-learning. The usability of the features of both interfaces will be evaluated. An e-assessment, experimental condition with two interface versions was developed to serve as a basis for this investigation. The chosen topic for the e-assessment was one that is often used in the design of software systems in education. The study consisted of two groups of users (a control group and an experimental group) in which the usability and performance of the two groups, specifically in terms of efficiency, effectiveness, and user satisfaction, were compared.

## 3.2 Aims

In line with the overall aims of the study this experiment aims to achieve the following:

1. To examine the impact of individual modalities on the effectiveness of the e-assessment interface in terms of usability. Specifically, avatars, images and description text within an e-assessment framework.

2. To investigate the usability in terms of efficiency (time taken to complete tasks) and user performance (score). A post-experiment questionnaire is conducted to derive user satisfaction.

3. To investigate the implications of varying degrees of difficulty in the questions for each metaphor and how this has an impact on overall usability.

4. To conduct a control experiment using normal text only which represents a traditional assessment interface without multimodal? The control experiment also tested for usability aspects, chiefly, user satisfaction, effectiveness and efficiency.

**3.3 Hypothesis**

Several usability criteria or standards have been improved to help designers to create a usable e-assessment system which provides learning content as opposed to just a confusing interface. This was carried out through the addition of description text, images and an avatar as components of a multimodal-based interaction with the e-assessment interface. Consequently, the hypotheses were formulated regarding the aspects of success of e-assessment:

• Ease of answer: how the online assessment is enhanced for users.

• Efficiency of use: how quickly students can locate the answer through the use of multimodal (using the description text, images and an avatar).

• Effective test: how effective the use of multimodal are as quantified by the number of questions successfully answered.

• Satisfaction: the student should enjoy achieving a given mission and will be more satisfied when the e-assessment is more usable.

In order to measure the learnability of VOAP and VMAP, a set of 4 hypotheses has been stated in order to compare the metrics in this experiment in terms of efficiency, effectiveness and satisfaction. These hypotheses are:

**H1:** The VMAP interface will be more effective than the VOAP in terms of the time spent by users to complete the tasks.

**H2:** The VMAP interface will be more effective than the VOAP as the task complexity increases.

**H3:** The VMAP interface will be more effective than the VOAP in terms of the provision of correct answers.

**H4**: The users of the VMAP interface will be more satisfied than the VOAP users.

## 3.4 Design of Experimental Condition

To verify the suggested hypotheses, two interfaces have been designed and advanced according to guidelines. The guidelines pertaining to the design of multimodal information, performance, [114] and multimodal user interface [115] were followed. An e-assessment condition was developed particularly to be used in conducting this experimental investigation. The interfaces provided two different condition versions: an assessment 'text only interface' version, and an assessment with multimodal.

Both condition versions of the empirical interface were designed to provide similar information about the test. This was represented in the form of two types of questions: true or false questions and multiple choice questions which included 3 difficult questions and 3 moderate questions as well as 3 easy questions. Each interface was presented in a divided screen display.

Table 1 shows the difference between the assessment presented and the interaction multimodal incorporated in one version of the experimental e-assessment condition. It can be noted that the VOAP use text only in communicating every type of task. Furthermore, the presentation of the assessment in the VMAP (images, avatar, and description text) was focussed on a multimodal approach in which diverse interactions were used to assist the users by explaining the questions in the e-assessment.

### 3.4.1 Non Multimodal Condition (VOAP)

Non-multimodal interface **(**VOAP**)** is the text only version. Essentially, it is the e-assessment interface with multimodal objects removed. Nevertheless, it still retains the general properties of the original condition with the same feel, assessment, order of chapters and level of questions.

### 3.4.2 Multimodal E-assessment Condition (VMAP)

VMAP is used in the assessment interface to introduce the avatar, description texts and images as added multimodal, displayed simultaneously in the interface. These avatar's expressions were specifically chosen based on the expressions typically used in daily life to express human feelings [116]. The plan is to use the facially expressive life-like avatar to narrate the explanation of small pieces of information for questions together with an interesting video, where the user may move the mouse cursor over the question and, following that, move the mouse cursor over the available answers from the multiple choices. The avatar will occupy the right side of the screen, so as to suggest to the user that it might assist them in selecting a correct answer and encourages the user to move the mouse cursor over the button to answer the question.

Every page's design has incorporated the feature by dividing the screen carefully, question by question, in order to avoid overlapping of questions, so the user can easily select answers on the screen. The left part shows a text of the question on a blue background, with a font size of 18 for the test. The avatar occupied the right side of the screen on a black background. When placing the mouse cursor on play in the video, the facially expressive avatar started to speak about and explain the question.

In addition to this, the right part of the screen shows the multiple choices available and allows the user to click the correct button. After the user had finished reading and listening to the information, and the user had selected an answer they thought was correct, the user moves the mouse cursor over to the button for the next task.

| Conditions | Features | | | | |
|---|---|---|---|---|---|
| | Sound | Description Text | Avatar | Images | Normal Assessment |
| M | √ | √ | √ | √ | |
| NM | | | | | √ |

**Table 3.1: Multimodal used to communicate the information by users in the multimodal interface group to answer the questions correctly**

The modality used in the test interface to introduce the question helps to explain the question to assist the user in selecting the right answer and helps them to save time. The left part shows a text question on the blue background, whereas the images occupy the right side of the screen. The images explain information about the question. In addition, the right part shows the various answers available, whether they be multiple choice or true and false, and permit the user to click the button. After the user has finished reading the question, been shown the modality and the available answers, the user takes a decision as to what they believe is the right answer and continue on to the next task by moving the mouse cursor over the button.

The next modality used in the assessment interface, namely, description test, to introduce the question, helps to explain the question to assist the user in selecting the right answer and helps them to save time. The left part shows a text question on a blue background, with font size 18 for the assessment. The description text occupies the right side of the screen on a purple background. In addition, the right part shows the answer either multiple-choice or true and false.

After the user has finished reading the question, been shown the modality and the available answers, the user takes a decision as to what they believe is the right answer and continue on to the next task by moving the mouse cursor over the button. Following all questions, the user is shown their score.

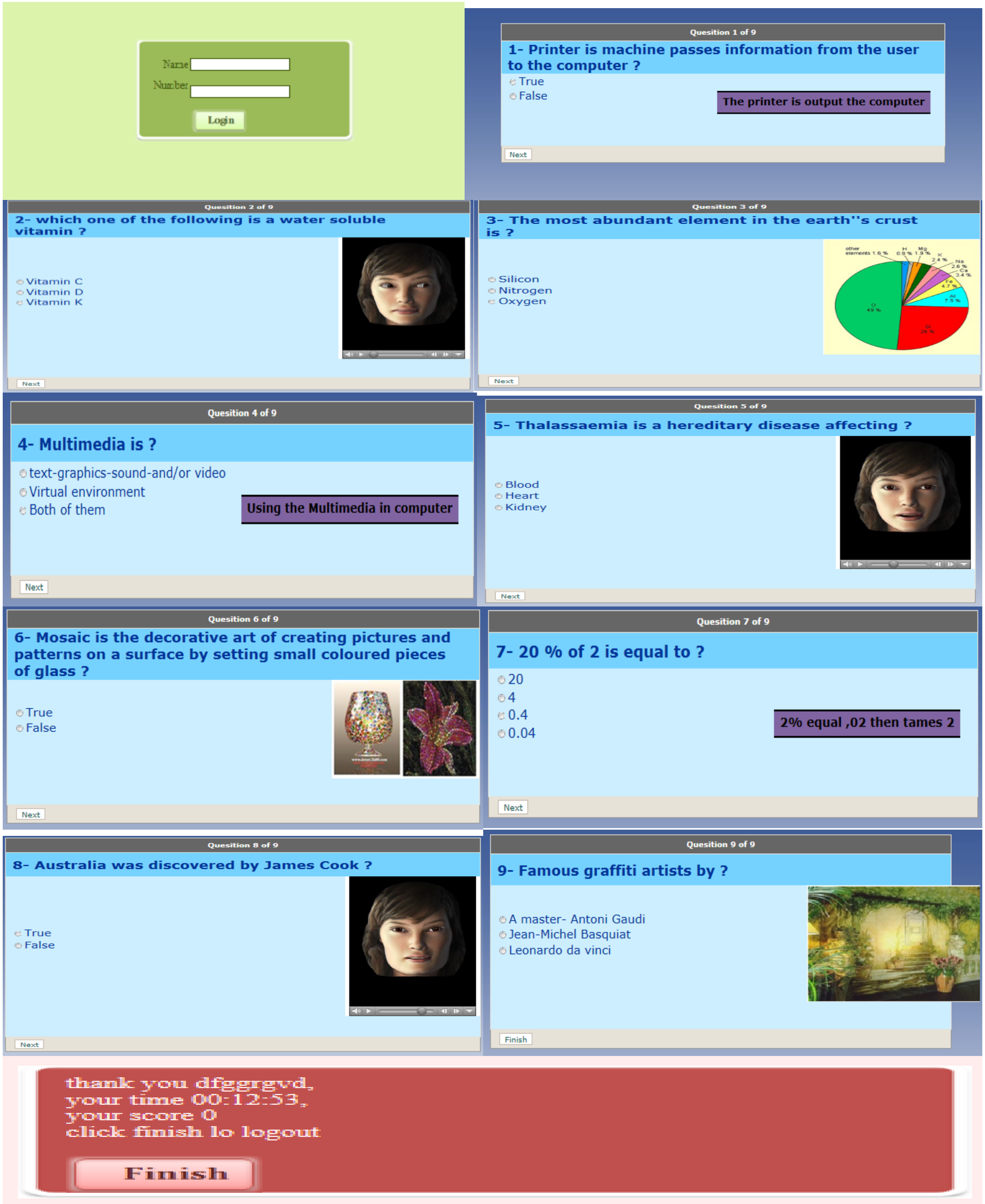



**Figure 3.1: Snapshots of experimental condition using the multimodal in e-assessment interface (VMAP) [110]**

### 3.4.3 Implementation of Avatars

The expressions of the avatars were chosen based on the expressions typically used in daily life to express human feelings and emotions [117]. In order to develop avatar presentations, the following tools were utilised:

 Audacity; the avatar is a different multimodal condition, ingredient that collects information from both auditory and visual human senses. It was used as an avatar, to record live speech sound and creates it in WAV (Waveform) data coordination.

 Mimic; here the avatar most resembles a human as cartoon-like characters [223]. Moreover, this software routinely includes mouth movements, eye blinks also head nods to personify the figures.

As well as this, it has been confirmed via numerous studies that the utilisation of avatars contributed definitely in terms of facilitating the learning procedure and enhancing users' attitude about online courses [118, 119].

### 3.4.4 Implementation of Images

Nowadays, visual stimuli are utilised to sell products, or to demonstrate the facts. Tests have proven that the proposed characteristic adoption strategies really enhance the precision, recall average in image retrieval. In addition, usability experiments which realise the effects of multimodal condition in information retrieval yielded that users really understand multimodality [120]. Colour images were used in this study to explain the questions to users through the information delivered in order to demonstrate the different aspects of the images.

### 3.4.5 Implementation of Description Text

This condition, as seen in Figure 3.1, included text and description text (VMAP). It therefore does not need command buttons as the text is presented in a clear manner with the information located near the question in a textbox. The question is explained to the user through the provision of information about the correct answer in e-assessment, thereby assisting the user in completing the e-assessment.

### 3.4.6 Design of Tasks

Each condition Figure 3.1 includes nine questions and each page contains one question. The task was divided into 2 types consisting of multiple-choice or true and false questions. Additionally the questions were arranged in terms of the levels of complexity (easy/moderate/difficult). This was conducted in order to investigate the impact of multimodal metaphors in condition and to identify which interface would be enhanced in terms of effectiveness, efficiency, and user satisfaction for the e-assessment procedure. One group tested the 'text only' interface, acting as a control group, and the other group tested the multimodal interface condition as an experimental group.

The design methods allowed users to be tested according to various experimental conditions, in order to determine if the effects were as a result of the individual conditions [121]. This design methodology, between-subjects testing, involves the assignment of different users to test different experimental conditions and therefore, guarantees controlling the learning effect [121]. Overall, 30 users contribute individually in the experiments and equally to both groups.

| **Users** | **E-assessment** (VOAP) | **E-assessment** (VMAP) |
|---|---|---|
| 15 | T1, T2, T3, T4, T5, T6, T7, T8, T9 | T1, T2, T3, T4, T5, T6, T7, T8, T9 |

**Table 3.2: Tack system**

| **Tasks Distribution** | **Question Type** | | |
|---|---|---|---|
| **Complexity level** | **Easy** | **Moderate** | **Difficult** |
| **Tasks** | T 3 | T 2 | T1 |
| | T 6 | T 5 | T 4 |
| | T 9 | T 9 | T 7 |

**Table 3.3: Tasks complexity levels**

## 3.5 Variables

The variables identified in the experimental design can be classified into three types: dependent, independent and controlled and are described in the following sections.

### 3.5.1 Independent Variables

Independent variables represent the aspects important in the experiment and are expected to be the reason of the results. The independent variable in this experiment was the test condition which is the role of the avatar with human-like expressions, text and also images. These variables included three levels of complexity in questions for the e-assessment interfaces; namely, easy, moderate and complex.

### 3.5.2 Dependent Variable

The dependent variable in this study was used to measure efficiency, by recording the time spent to complete tasks by each user; and effectiveness, by the correctness of answers to tasks; and monitored responses to subject satisfaction using the Likert criteria in both conditions.

### 3.5.3 Controlled Variable

The variables expected to influence the experimental process were controlled. The variables in this experiment were:

1. Tasks: every user was knowledgeable of the same number of tasks in all of the interface conditions.

2. Subject: the subject tested in the experiment was similar in both interface versions, with questions about the same level of difficulty.

3. Consciousness of tasks: users were conscious of the tasks that would be provided to them.

4. Time: users had sufficient time to complete all particular tasks in both interfaces. Consequently, a task completed within the time allocated would be regarded as successful or else the task would be envisaged as unsuccessful.

5. Condition familiarity: each user was given the same instructions and it was the first time that each user tested the interface or condition.

6. Consistency: The execution of the experimental interfaces execution was examined among the similar users on a personal basis to all users.

Furthermore, the same process was followed during the procedure of the experiment whilst using the same computer for demonstrations.

## 3.6 Data Gathering and Questionnaire

Both user groups were made to follow the same procedure in a bid to keep the experiment consistent. Both experimental and post-experimental elements were components of the nine-page-long questionnaire (see Appendix A-1). An introduction to the study can be found on page 1, outlining the experiment’s aim as well as informative steps as to how the questionnaire should be completed along with an explanation of the five-point Likert scale. Page 2 allowed users to enter information such as gender, age, internet and computer experience, education level as well as previous knowledge of e-learning and avatars. The questionnaire’s post-experimental element assisted in determining user satisfaction under the e-assessment conditions tested.

The satisfaction score was calculated through user responses, both for the experimental and control groups. The tasks testing the second set of conditions were placed thereafter, structured in a similar manner to the first set of tasks.

### 3.7 Pilot Study

A preliminary pilot study was arranged and conducted on a small scale with six users in order to achieve the following:

1. Evaluate whether the questionnaires were simple to understand.

2. Evaluate feasibility and whether the interfaces and the instructions were simple to recognise.

3. Validate the experimental process and assessment, as well as the desired typical time to complete the tasks.

4. Determine some other usability issues that were not already picked up through the design model which could then be adjusted.

5. Gather some other feedback that the pilot users provided so that it could be consider in experimental design.

6. The pilot study helped in the experiment to control time.

7. The pilot study helped in the experiment by showing levels of the question from easy to difficult.

### 3.8 Users Sampling

Overall, 30 users participated in this thesis independently. The users were equals and randomly assigned; (N = 15) to the experimental conditions; text with e-assessment interface for the control group, and a multimodal interface, (N = 15) for the experimental group in the library of De Montfort University. All students used both e-assessments: VMAP and VOAP. They were distributed randomly and equally for each user, VOAP to the experimental conditions; e-assessment interfaces without multimodal for the control group, and multimodal condition VMAP for the experimental group. Moreover, the sample was chosen randomly, and there was no agreement or proposed time.

Also, the selected users were selected within the university library and the bulks were unaware of what they were being asked to do. Most of the users in each group had some prior experience of e-assessment, indicating that they will may rely on the communicated information to answer the required task.

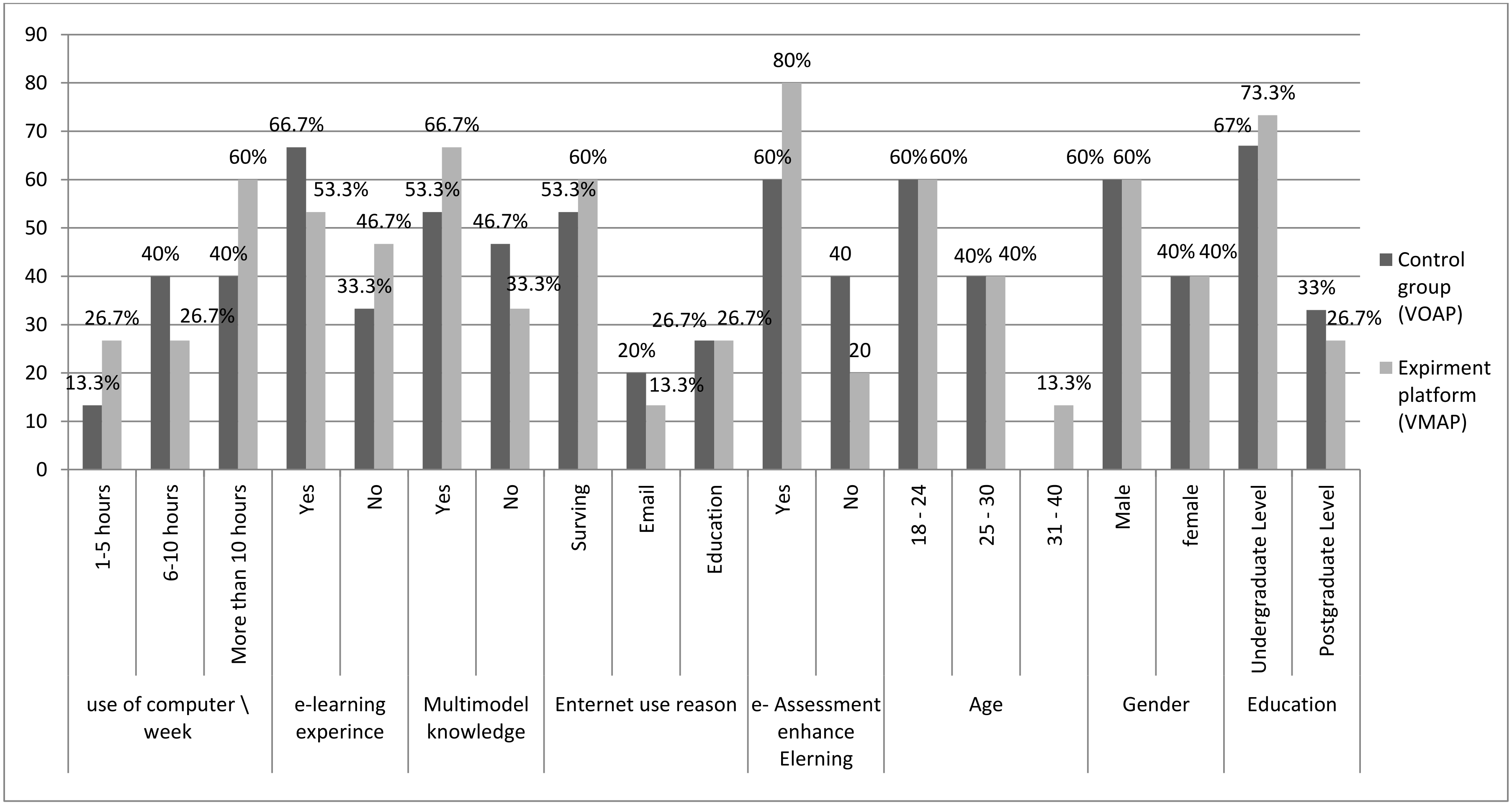



**Figure 3.2: Users profile in terms of age, gender, education level and area of study and prior experience for users in both groups**

## 3.9 Results and Analysis

### 3.9.1 Users Profiling

The data in relation to users' profiling as well as educational information, as shown in Figure 3.2, demonstrates that the age range in the control group was varied, with around 60% within the 18– 24 demographic and 40% consisting of 25 - 30 years old; 60% of students who participated in this study were males (9 altogether) and 40% (6 altogether) were female; these constituted members of the control group. In the experimental group, the ages were 80% within 18 – 24, 20% within 25 – 34, and 0.0% 31 – 40 years of age. The education stage was found to be predominantly undergraduates with 67 % present in the control group, the remaining 33.3% being postgraduates. In the experimental group, 73.3% of the participants were undergraduates with 26.7% taking part as postgraduates. Also, as can be noted in Figure 3.2, 13.3% of the control group reported to use computers for 1-5 hours a week, and 40% reported to use them between 6 – 10 hours a week. Furthermore, 46.7% of participants said they used computers for more than ten hours a week and 26.7% for 1-5 hours and a further 26.7% between 6 – 10 hours per week. The remaining 46.7% reported to use computers for more than ten hours a week. With respect to e-learning system experience, around 66% of the control group had more experience compared with 46.7% who had some experience with HCI. It was also found that the most popular reason for internet use was surfing with 53.3% of the control group claiming this was the main usage of the internet and a further 60% of the experimental group claiming the same.

Furthermore, more than 66.7% and 46.7% had good multimodal knowledge in both groups, respectively, and at the very least, users had a limited background in multimodal knowledge in both groups. Additionally, Figure 3.2 demonstrates that from the experimental group, the majority of the users (80%) said yes: e-assessment is a good way to enhance e-learning applications, in comparison with the control group where 40% disagreed.

### 3.9.2 Experiment Sessions

The test took between 3 and 6 minutes with a mean time of 5.46 minutes, not including post-experimental and pre-experimental questionnaires. The time was distributed as follows: Users engaged with the pre-test questionnaire for around 3 minutes and read the tasks from the questionnaire for about 4 minutes.

The users started with the first interface, for example, VOAP, for around 10 minutes. Upon completion, students filled out the five-point Likert scale presented in the questionnaire which took about 4 minutes. This process was similar for the second interface. In conclusion, the user would finish the experiment along with the post-questionnaire: all tasks taking approximately 4 minutes.

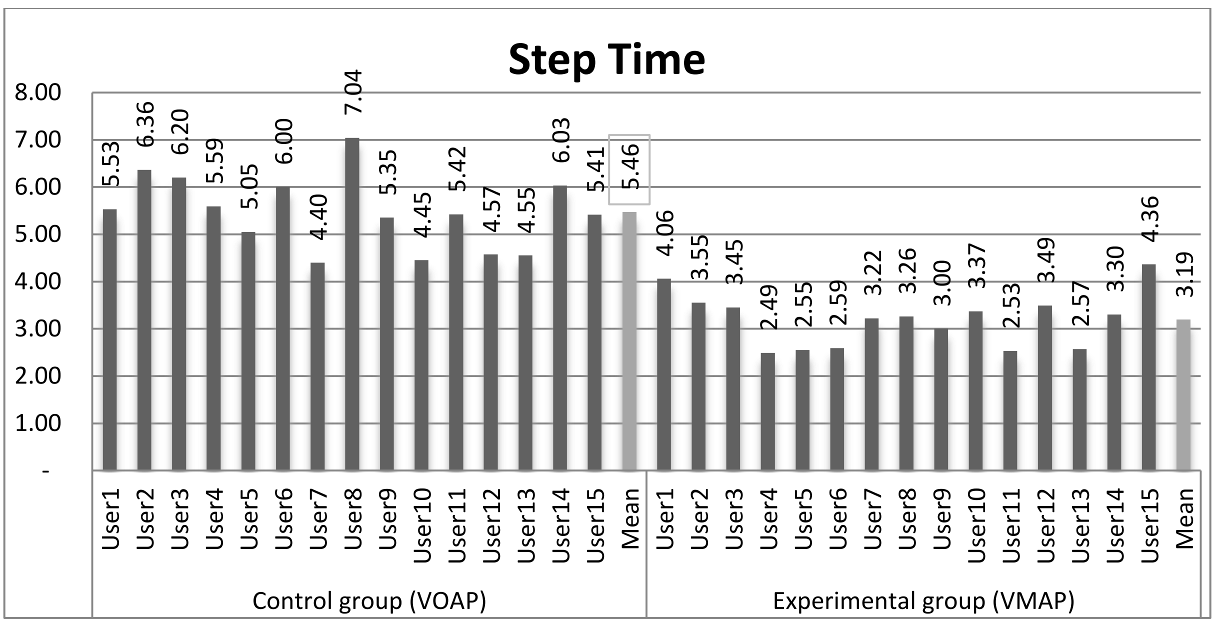



**Figure 3.3: Total time taken by each user in both groups to answer every question**

### 3.9.3. Individual User

Figure 3.3 shows the overall time spent by every user in all groups to answer all of the nine questions. Users of VOAP spent more time on the questions compared to users of the VMAP. However, the observed time differences between tasks are highly significant, with the lowest and highest recorded times for the first condition control group being 4.40 minutes (User 7) and 7.04 minutes (User 8) respectively; the mean time being 5.46 minutes. In the second condition experimental group, the maximum time recorded was slightly lower (2.49 minutes by User 4) and the minimum time (4.36 minutes by user 15). The mean time recorded was 3.22 minutes. In short, using multimodal metaphors in communicating the e-assessment material enabled the users in the experimental group to outperform their counterparts in the control group in time spent answering the required questions. (See Appendix A3 and A4).

### 3.9.4 Efficiency

The diagram 3.4 shows the mean value of the time taken by users to answer the questions this was used as a measure of efficiency. This measure was considered for both VOAP and VMAP groups, for all tasks, in accordance with the question complexity (moderate/easy/complex), for all questions and all users in both control and experimental groups. It can be seen that overall, the time taken to answer questions was shorter in the VMAP group. Experimental observations showed that users in the control group regularly divided their visual attention between the symbols provided, which indicated assessment code and assessment content to understand the presented information, and in some cases a visual overload may have occurred. The users in the experimental group, however, kept their visual attention directed towards the assessment content. The raw data for all questions, answering time can be found in Appendix A-3 and A-4.

#### 3.9.4.1 Levels

Figure 3.4 explains the answering time grouped by the complexity of the questions which were designed to increase in difficulty and were equally divided into 3 categories: easy, moderate and difficult. All users had to answer nine questions in total. Figure 3.5 illustrates the mean time taken to answer all questions using the VMAP condition. This did not include reading time as well as the time taken to fill out the pre-task and post-task questionnaire.

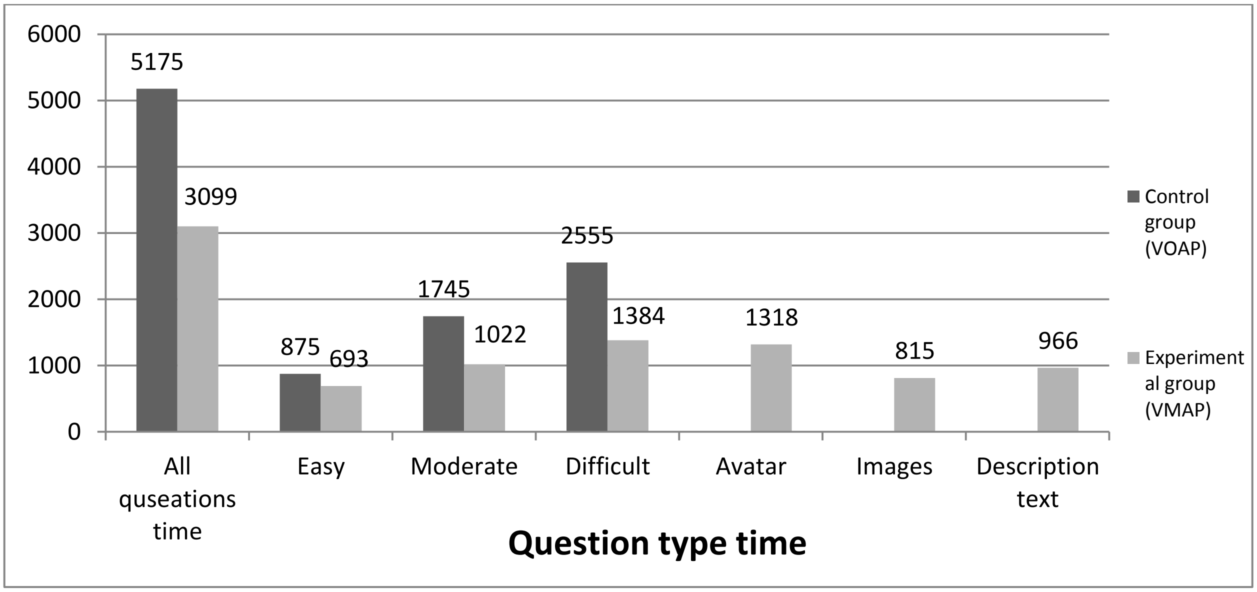



**Figure 3.4: Mean values of time taken by users in both groups to answer all and grouped through question complexity**

Overall, the total time taken recorded by users of the VOAP in the control group was 5175 seconds, averaging 5.46 minutes for each user, compared with users of the VMAP in the experimental group who took a total of 3099 seconds to answer questions, averaging 3.19 minutes for each user. It can be observed that users of the VMAP were 2076 seconds faster than those who used the VOAP. The t-test calculations illustrate that the difference between both groups in answering time was significant ($t = 32.252$, $MD = 345$, $sig < 0.5$). The experiments revealed that users in the VOAP group directed their vision towards the questions located in the text box. However, users in the VMAP group maintained their visual awareness to the images and description text, though they were listening to the avatar’s messages which helped them to enhance their focus on the delivered answer ($t = 25.246$, $MD= 206$ $sig < 0.5$).

To summarise, the users in the VMAP group were considerably aided through the addition of the multimodal metaphors in the experimental group, which enabled them to spend less time, compared with the users of the control group, in answering the questions in the e-assessment. Overall, it is observed that the answering time in the experimental group was lower for all complexity levels. It can be seen that the difference in task completion time between the two groups increased as the level of task complexity increased, except in the moderate and complex level questions of the experimental group.

Figure 3.4 illustrates the total time spent by each user in the experimental group VMAP to answer all questions. It demonstrates the result of each multimodal implementation (avatar, images, description text), which all included 3 questions. Users of the images multimodal observed a time which was slightly shorter (815 seconds) when compared with users of description text and avatar multimodal in the experimental group. The second shortest answering time was by users of the description text multimodal who took an average of 966 seconds, the difference between images and description text being 151 seconds. Finally, users of the avatar multimodal took the longest (1318 seconds) to answer the questions. In short, users of the images multimodal using VMAP were 815 seconds quicker compared to users using description text and the avatar multimodal to assist in answering the questions.

Figure 3.5 illustrates the total time spent by each user in the experimental group VMAP to answer all questions. It demonstrates the result of each multimodal implementation (avatar, images, description text), which all included 3 questions. Users of the images multimodal observed a time, which was slightly shorter (815 seconds) when compared with users of the description text and avatar multimodal in the same experimental group. The second shortest answering time was by users of the description text multimodal who took, on average 966 seconds, the difference between images and description text being 151 seconds.

Finally, users of the avatar multimodal took the longest (1318 seconds) to answer the questions. In short, users of the images multimodal using VMAP were 815 seconds quicker than users using description text and avatar multimodal to assist in answering the questions.

### 3.9.5 Effectiveness

The percentage of correctly answered questions was used as a measure of effectiveness. This measure was considered for all the tasks in total, according to the multimodal type and question complexity (easy, moderate and difficult) as well as for each user in both control and experimental groups.

#### 3.9.5.1 All tasks

Figure 3.5 presents the variation between users' performance in relation to the different multimodal, namely; avatar, images, and description text, in terms of correct answers provided by users. In consideration of this, for those who used images, 80% gave correct answers, compared with a 75% for those who used the avatar multimodal and 71% for those who used description text (Figure 3.6).

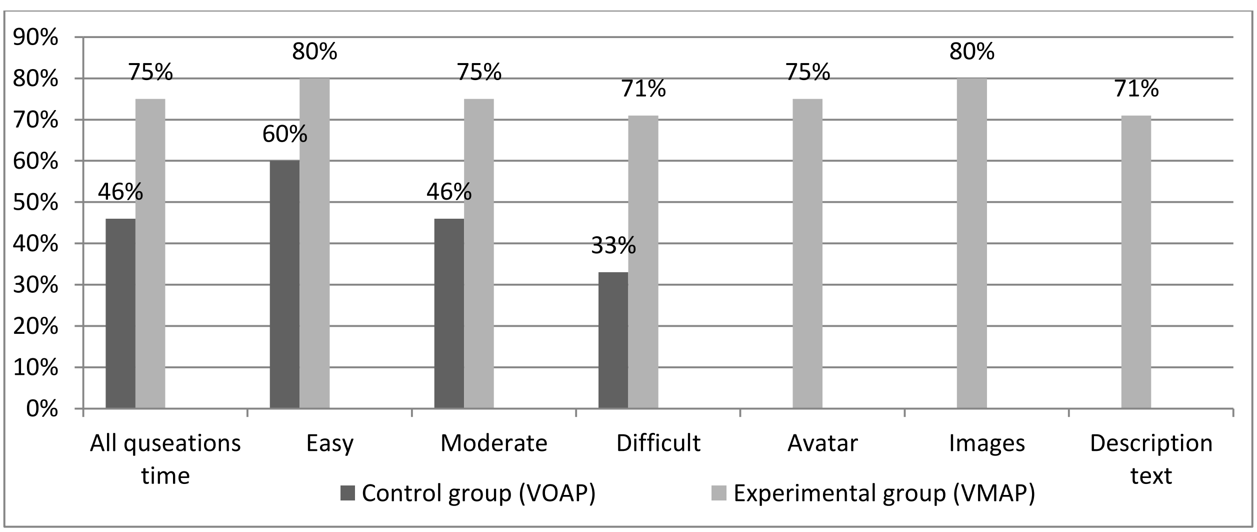



**Figure 3.5: Values of percentage correct answers from all questions taken by users in the VOAP and VMAP by complexity level**

Correctly answered questions were used to measure the effectiveness of the metaphors. Figure 3.5 presents the correct answers, as a percentage, for questions in both the VOAP and VMAP. For the latter 75% answered correctly, more than in the VOAP at 46%.

The total percentage of correct answers given by users of the images multimodal were 36% indicating that users find information, communication via images to be easier than via avatars and description text. Subsequently, users of the avatar multimodal achieved a 33% score with the lowest percentage of correct answers given by users of the description text multimodal at 32%, making it harder to complete. The mean percentage of correctly answered questions in the VMAP and VOAP was 7 and 5 respectively.

As shown in figure 3.5, the integration of more than one communication metaphor of a different nature in the VMAP model assisted users in the experimental group in highlighting the different kinds of information which had been delivered through each of the metaphors (description text, avatar, images). As a result, they outperformed the users of the VOAP who received the learning information by images firstly, followed by an avatar and finally by description text. Conclusively, the multimodal interaction metaphors used in the VMAP were more effective in communicating the learning material and considerably assisted the users in the experimental group to achieve a more effective rate, as opposed to the control group users.

Figure 3.5 illustrates the percentage of correctly answered questions for all levels of complexity, namely easy, moderate and difficult questions for both groups. The results show that the control group was outperformed by the experimental group; this was particularly noticeable for difficult questions. What is more, for easy questions in the VMAP a score of 80% was achieved, more than that about the VOAP condition. However, a larger difference between the two groups was observed for moderate questions and the largest difference was noted for difficult questions. For the VMAP condition, the users in the experimental group answered questions correctly with a score of 80%, 75% and 71% for easy, moderate and difficult questions respectively. In contrast, the users of the VOAP produced a score of 60%, 46% and 33% for the easy, moderate and difficult questions respectively. In summary, it is clear that for easy questions both groups of users achieved equivalent levels of accuracy. However, multimodal metaphors contributed significantly more to a better score with higher complexity questions.

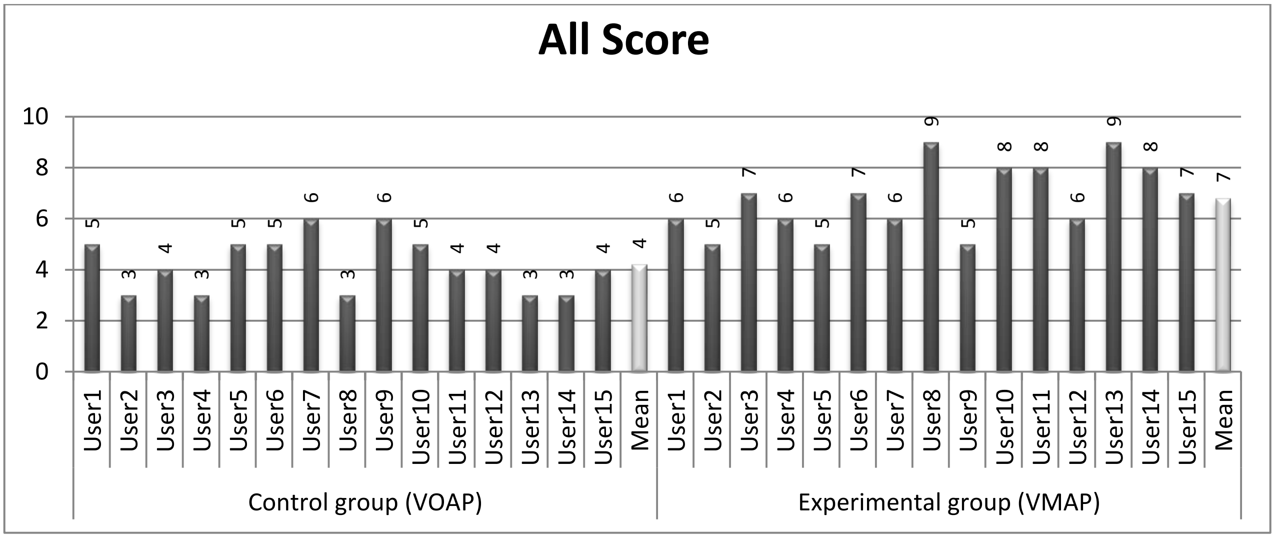



**Figure 3.6: Mean total number of correct answers by each user in condition**

### 3.9.5.2 Individual User

Figure 3.6 illustrates the total number of correct answers given by each user in both groups: VMAP and VOAP. It is commendable to note that 2 users (8 and 13) of the experimental group correctly answered all nine questions, with a further users (10, 11 and 14) answering all but one (question 8) correctly.

However, none of the control group users were able to reach this level of performance. In fact, the highest achievement recorded was 4 correct answers given by users 7 and 9. In addition to this, the users with the lowest score in the VMAP group (2, 5 and 9) were able to correctly answer 5 questions, whereas the users with the lowest score from the VOAP group (2, 4, 8, 13 and 14) were only able to correctly answer 3 questions. The mean, as shown in Figure 3.6 was higher in the VMAP group (9) compared to the VOAP group (6). In summary, the use of multimodal in communicating the information enabled the users of VMAP to outperform the users of VOAP in answering the questions correctly. The raw data for all questions, answering correctly, can be found in Appendix A-5 and A-6.

### 3.10 Satisfactions

Users were able to express their attitudes to pre-selected statements via the post-experimental element of the questionnaire using the five-point Likert scale, thereby enabling the measurement and recording of user satisfaction. The pre-selected statements were associated with ease of use, complexity, confidence, ease of learning, and also general satisfaction. Each statement was scored using the five-point Likert scale with options ranging from 1 (strong disagreement) to 5 (strong agreement). When calculating the overall satisfaction score, this was conducted using the SUS (System Usability Scale) system [122]. The scoring method followed here is to take the average score for each statement. This mostly results in a positive impact where users liked the VMAP condition more than VOAP condition.

Statistically, the t-test proved that the difference in users satisfaction among both groups was significant ($U = 50$, $CV = 72$, $p < 0.05$). In other words, the VMAP was more satisfactory than the VOAP. Figure 3.7 shows the frequency of the user agreement with every statement in the satisfaction questionnaire. High levels of agreement were shown via the users in control and experimental groups for difficulty using the system (S8). Nevertheless, the VMAP was less time while interacting with the system where I felt worried (S4), and high levels of agreement could be easier to learn (S7) as opposed to the VMAP.

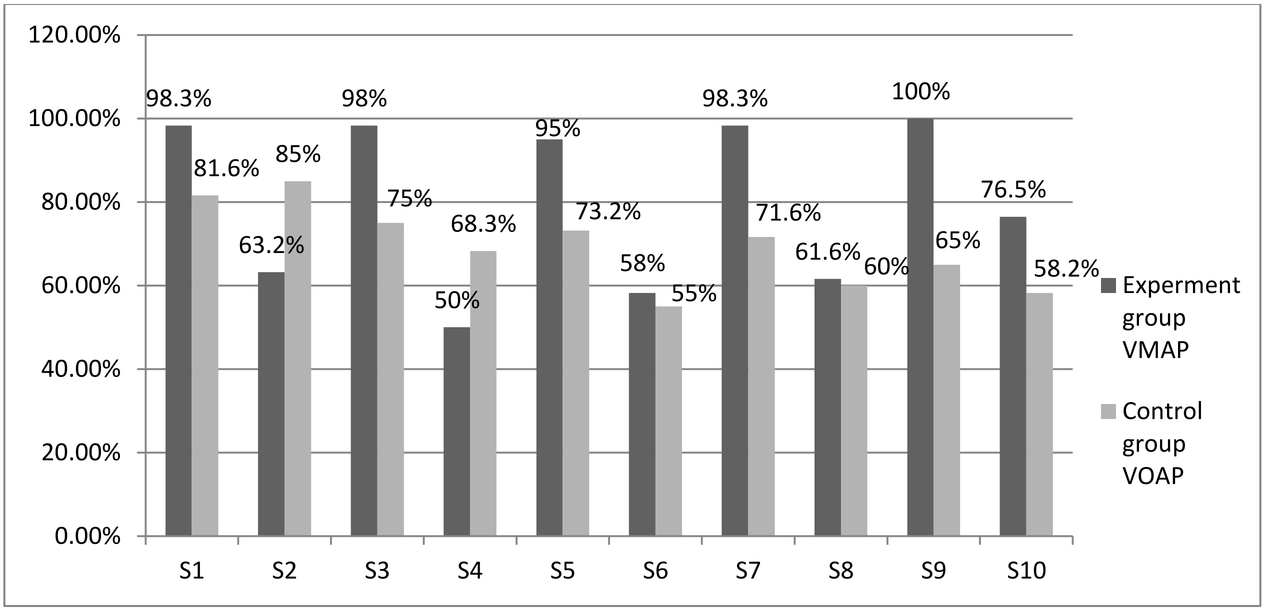



**Figure 3.7: Frequency of users' agreement to each satisfaction statement in both VOAP and VMAP condition**

In the first statement (S1), 98% and 81% of users in both groups agreed that the tested e-assessment interfaces were easy to use. The statement (S2) asked the users whether they discovered the system to be unnecessarily complex. Users of the VMAP show a slightly higher level of disagreement 63.2% than the users of the VOAP 85%. It can be noticed that 98% of users for S3 in VMAP thought was easy to use the system, but in VOAP this was 75%. In relation to (S5), the entire VMAP 95% of users found that all functions were well integrated compared with 75% for VOAP. The users found the VMAP condition was easy to understand where the mean satisfaction in the experimental groups was 79% compared to 69% for the users in the control group.  The users' satisfaction is significantly enhanced in the VMAP interface in comparison with the VOAP. The raw data for satisfaction can be found in Appendix A-7.

### 3.11 Discussion

The experiment aimed to determine the best use of metaphors to enhance the performance of e-assessment and to measure the efficiency of the two conditions for every task by measuring the time spent and correct answers. The obtained results have been used to evaluate the two conditions in terms of efficiency, effectiveness and user satisfaction. In addition, the experiment is based on factors that might evaluate the role of multimodal interaction, for example complexity stage, (easy, moderate and complex). Consequently, these results are discussed from the following three measures through the multimodal in users ‘efficiency, effectiveness and satisfaction.

1. Time taken to answer the questions in e-assessment in terms of complexity levels and efficiency.

2. Correctness of users, answers to the questions with respect to the use of avatar, images, description text (effectiveness of multimodal performance) in the experimental group.

3. User satisfaction - with both of the e-assessment condition.

The experiment (VMAP) produced encouraging outcomes and the results illustrate that the use of multimodality (avatar, images, description text) was more efficient and effective also more satisfactory compared to using the text without multimodality. The hypothesis assumed that the multimodal e-assessment condition will be more efficient than the text without multimodal in relation to the efficiency of users in answering the questions. The experimental (VMAP) results, as shown in Figure 3.3, a show the differences between interfaces are not significant, and demonstrates that using the multimodal interaction decreases the total time needed by users. The insertion of dissimilar multimodal communications in the experiment group assisted the users to focus on the presented information during the auditory channel, while at the same time using the visual channel to know this information [123].

Consequently, users were significantly assisted through the adding of these multimodal in the VMAP in terms of consumption reduction and answering time compared with users of the VOAP interface. Moreover, these results recommended that using the images, avatar and description text could be significantly more efficient than using the text without multimodality in presenting clarification concerning the questions in this experiment. With regards to complexity, it was hypothesized that the VMAP will be more effective compared to the VOAP with an increasing level of difficulty of the questions. Therefore, the results of this experiment in both conditions (refer to Figure 3.6) illustrate an increasing difference in answering scores and efficiency which support H3. The VMAP outperformed the VOAP interfaces in answering easy questions, which showed the beneficial influence of multimodality. A larger contribution of the multimodality was that users in the VMAP achieved significantly higher level of correct answers compared with the VOAP for questions that were moderate to difficult. This result provided evidence of the gradual impact of multimodality with an increasing level of complexity and also demonstrates that users' performance can be enhanced by the use of these multimodal in e-assessment condition, thus, these results proved the assumptions made in H2.

Considering the experimental results, as expected in H3, the users of the VMAP achieved a larger number of correct answers and less time to complete the obligatory tasks than the VOAP users. Information in the VMAP was provided as images and not movement, while other aspects of the multimodal were presented as an avatar and description text. The results of this experiment indicated that images enabled users to learn better and easily in terms of time and correct answers (see Figures 3.4 and 3.3). Overall, it was anticipated that users of the VMAP could be more satisfying than the users of the VOAP. Consistent with this assumption, the multimodal presentation of the e-assessment condition in the VMAP has proved superior to the text without multimodality in the VOAP. Consequently, the user expressed a more helpful attitude towards images in e-assessment condition.

Nevertheless, both of the tested e-assessment interfaces were easy to use and provided ease of learning, and neither was confusing the obtained results did not confirm a significant difference between both interfaces in terms of satisfaction. In fact, users in the VMAP retained information from e-assessment interface for less time than in the control group, enabling most of these users to achieve a higher number of completed answers compare to VOAP.

These findings proved the assumptions in H3. Therefore, the user's ability to assimilate and understand the information provided through multimodality could furthermore be enhanced e-assessment condition. This outcome was derived from two dependent and independent groups and users within those groups were offered both condition versions in order to make a comparison. Users in the experimental group expected that their learning was improved, aided through the multimodal metaphors. It was easier for users to identify the learning information in the assessment, which has been facilitated by images, avatars and description text. When user satisfaction, effectiveness and efficiency outcomes are combined, the argument that users in the VMAP were assisted by multimodality becomes stronger. It can conclude that the use of multimodality helped in e-assessment condition is more likely to result in an enjoyed and greater satisfying experience for the user.

This experience is related to the ability to achieve learning tasks correctly. In brief, the results of this experiment suggest the importance of integrating multimodality in e-assessment condition, especially images, to improve learning performance and usability in terms of efficiency, effectiveness and user satisfaction.

The experimental e-assessment interface will be further developed in order to test the hypothesis of this research and to answer the research question and finally to achieve an aim of this research which is to generate guidelines for the inclusion of multimodality in e-assessment interfaces. The addition of multimodality such as avatars with body gestures could also be employed. A significant number of comparative usability experiments can be carried out in order to test multimodal features in the interface of the e-assessment, and to find out the role of each type of metaphor and which give the best results.

### 3.12 Summary

This chapter has presented the empirical experiments for investigating the role of e-assessment interfaces, and to evaluate multimodal metaphors (avatar, images, description text) on usability (in terms of efficiency, effectiveness and user satisfaction) and learning performance in e-assessment of Human-Computer Interaction. The major objective was to enhance the learning process during the development of two different versions of the experimental e-assessment interfaces. The first version, VOAP, was based on text without multimodality in presenting the task assessment.

On the other hand, the second version, VMAP, employed multimodal metaphors (avatar, images and description text) to deliver information. Both e-assessment interfaces were evaluated via two independent groups of users. The first group control VOAP tested text without multimodal condition and the second experiment VMAP tested the multimodality in terms of how it helped in achieving tasks and answering questions. The results of the experiment presented in this chapter indicated that the multimodal metaphors could indeed assist in the enhancement of the usability of e-assessment interfaces by reducing the time taken and enabling users to find the correct answers, thereby facilitating the assessment and increasing user enjoyment, as well as making the improved additions to interfaces more satisfactory.

In brief, it can be summarised that the multimodal metaphors which were tested could have supported the enhancement of users learning performance as well as the usability of the e-assessment condition of e-learning interfaces in terms of effectiveness, efficiency and user satisfaction. As a result of this, the next experiment (Chapter 4) was designed to identify and determine the usability (in terms of effectiveness, efficiency and user satisfaction) and learning performance pertaining to the utilisation of images with descriptive text, recorded speech and avatars in the e-assessment interfaces.

# Chapter 4

# Experiment II: The Role of Avatar in E-assessment

## 4.1 Introduction

Experimental results from the first experiment indicated that the use of multimodal metaphors has the potential to enhance user-performance and the importance of images, Avatar and the usability of the proposed e-assessment interface. Specifically, integration of Avatar, images, and description text, proved to be effective for assisting users in their learning. Experimental results shown from the first experiment, do not very obviously represent the role of all of the individual multimodal metaphors, as the objective of the study was to investigate the use of multimodal metaphors in terms of their effectiveness in assisting students during the e-assessment, user satisfaction and enhancing efficiency. Therefore, an investigation and comparison of the three new conditions, integrated with different aspects assisted students. This chapter describes the second experiment that has been conducted to discover and compare the role of avatars with images and naturally recorded speech with images and description text within e-assessment interfaces, to present three different questions. The question in the condition includes of two types of questions involving true, false and multiple choice questions, including 3 difficult questions and 3 moderately difficult questions as well as 3 easy questions.  Additionally, the next sections offer a detailed explanation of the study's aims and objectives, hypotheses, experimental interfaces, design of the experiment, outcome and discussion.

## 4.2 Aims

In line with the overall aims of the study this experiment aims to achieve the following:

1. To examine the impact of a combination of images with avatar, recorded speech and description text respectively on the effectiveness of the e-assessment framework.

2. To investigate the usability in terms of efficiency (time taken to complete tasks) and user performance (score). A post-experiment questionnaire is conducted to derive user satisfaction. Here a usability scale (SUS) is employed.

3. To investigate the implications of varying degrees of difficulty in the questions for each combination of metaphors and how this has an impact on overall usability.

4. To explore whether there were any significant differences with regards to effectiveness, efficiency and user satisfaction between the three interfaces that were tested.

5. To determine the improvement in learning performance of each metaphor combination.

### 4.3 Tasks

1. Formulate experimental hypotheses.

2. An improvement of three experimental conditions that use three different presentation modes, recorded speech with images, Avatar with images, and finally description text with images.

3. Experimental violation of the three conditions (recorded speech with images RI, Avatar with images AI, description text with images DI) using a study to examine and assess each condition. This comparative study will assist the evaluation of usability.

4. Measure efficiency in terms of time taken by users to complete each task in all interfaces.

5. Measure effectiveness through correct answers.

6. Measure user satisfaction for all experiments by allowing users to rate the three experimental interfaces against a set of criteria.

7. Analysis of the outcomes in a comparative study in order to determine the suitability of each of the metaphors used in the context of e-assessment.

## 4.4 Hypotheses

This section determines six hypotheses based on the inclusion of multimodal namely; images with avatar, recorded speech and description text and how they would impact the usability and learning performance in e-assessment condition. Based on this the following hypotheses were formulated:

H 1: A video using an avatar will be evaluated more positively by user when used in interactive e-assessment condition for communicating information.

H 2: The efficiency of images with avatar, recorded speech and description text will be dissimilar in terms of time spent to complete the same tasks by users.

**H 3:** The effectiveness of images with avatar, recorded speech and description text will be dissimilar in terms of number correct answers made and number of questions completed successfully.

**H 4:** There will be differences among the experimental interfaces in terms of user's satisfaction.

**H 5**: Avatar with images will be more efficient than description text with images and recorded speech with images in terms of shortening task achievement time.

**H 6**: Avatar with images will be more effective than recorded speech with images and description text with images in terms of reducing frequency of incorrect answers.

## 4. 5 Experimental Condition

Although the previous experimental interface proved to be successful, this section provides the e-assessment experimental interface used to empirically achieve a number of multimodal interaction metaphors proposed to enhance the usability of the e-assessment condition.

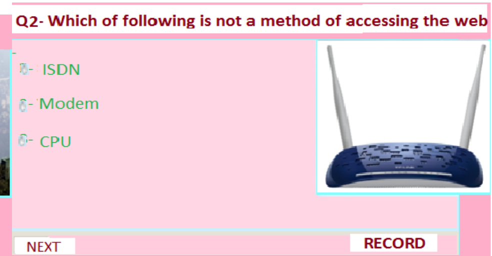
Three different e-assessment interfaces were designed to utilize speaking avatars and images as well as naturally recorded speech and description text in order to present audio-visual learning material. These conditions were (AI) speaking avatar with facial expressions with images, (RI) recorded speech with images and (DI) description text with images. It is thought that using avatars in this method imitates the traditional face-to-face interaction that typically takes place in human computer interaction. The implementation of AI, DI and RI condition involved similarities and differences.

### 4. 5. 1 First Interface Design (Record Speech and Images RI)

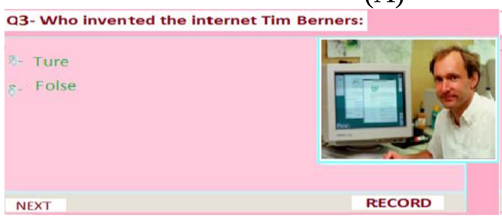
The approach used in this interface differs from the previous condition ; simply the user is shown images and listens to specific naturally recorded speech (see figure 4.1(A,B,C)) and must go directly to the task (questions) designed to be "general questions", the interface also provides pause / repetition connected to the delivered information that can be requested to get additional explanations from the naturally recorded speech to enable the user to think in order to help them extract the answers at any point and to enable users to interact with recorded speech and images only when they want to do so.

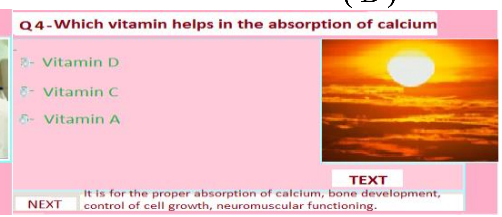
Each experiment is timed to show the time spent by the users for every task, and a particular button to stop the recorded speech if the user chooses not to continue or pause if the user gets the answer early. Time is limited for the users to listen to the recording to a minimum of 1-2 minutes. When the users have completed listening to the recorded speech seeing the images with color, it might assist the user in addition to the recorded speech. Additionally, this makes the likelihood of finding the correct answer higher and easier. Also, it encourages the users to get the answers from their first try; on other hand, without the multimodal the chance of finding the correct answers will decrease with the second and third attempts.

The user navigates to the right answer and chooses the true button or false button and the page automatically navigates to the following new interface until the user has finished all the questions.

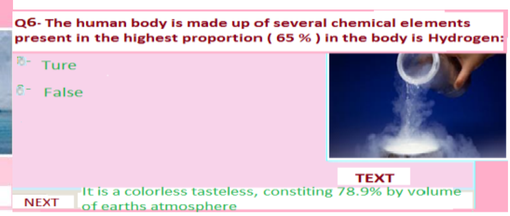




(A) ( B )

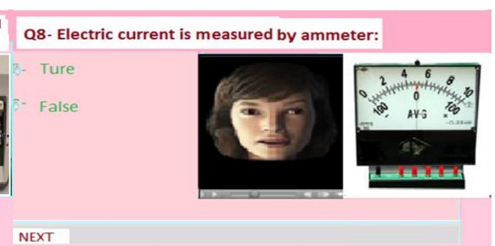



C (D)


(E) (F)


(G) (K)


(L)

**Figure 4.1: Snapshots of Experimental, Record speech and images A,B,C Interface. Description text and images D,E,F. Avatar and images G,K,L in E-assessment Condition**

### 4. 5. 2 Second Interface Design (Description Text and Images DI)

Description text with images in figure 4.1(D, E, F) has now become an effective method of information conveyance to users. In fact in this experiment description text with images was tested in an e-assessment interface, used to provide information to answer in an interactive way. The information is presented through a description text with the images acting as an extra multimodal. Directly under the description text is a clock showing the user the time spent on every task, when the users have completed the reading of the description text and see the images with color, it might assist the user in addition to the description text.

### 4.5. 3 Third Interface Design (Avatar and Images AI)

This Interface employed an avatar with images as can be seen in Figure 4.1 (G, K, L). The condition offers command buttons to enable the opening of the Avatar to be presented. The Avatar on the right hand side of the condition presents the information supported by a brief textual display in the condition to introduce the learning materials, as well as aspects to assist users. The condition provided pause / repetition functionalities related to the delivered information.

| **Conditions** | **Features** | | | |
|---|---|---|---|---|
| | Record speech | Description Text | Avatar | Images |
| M | √ | √ | √ | √ |

**Table 4.1: The condition multimodal features**

A similar procedure has to be followed by asking and answering questions in the next interface, in order to insure the consistency of the experiment, assessment questions were similar for all users. The system interacts with users and delivers information through an Avatar. Directly beneath the avatar, is a timer showing the user the time spent to complete each task, with a special button to stop the avatar if the user chooses not to continue or to pause the clock if the user gets the answer early. Time is limited for the users to listen for a minimum 1-2 minutes for every Avatar, measured according to the usual reading rate by the developer and according to the avatar's speech length. When the users have completed listening to the Avatar, the images with color may assist the user in addition to the Avatar, the user navigates to the right answer and chooses the button and the page automatically navigates to the following interface until the user completed all question.

### 4. 5. 4 Experiments Design

This design guarantees the involvement of all users in evaluation of all the systems being tested; thus, it reduces the impact of any other external factors that might affect user performance [121]. Furthermore, it requires all users to offer an effective contribution in the testing of each and every system. In order to fulfill this, one group of users participated in the testing of the experimental e-assessment interfaces, namely: AI, RI, and DI. Nine questions in the assessment were communicated using the experimental interface with 30 users participating in the experiment individually.

All the conditions were shown in a random order to avoid the impact of any other factors. All variables in this experimental study, including those of an independent, dependent and controlled nature were specified.

### 4.5.5 Experiment Tasks

The systems offer nine tasks (Questions) and all questions are multiple choice and true-false questions. These questions were created in the system, and every question is considered as one task. All three interfaces have variation in complexity level of questions, namely; easy, moderate and difficult. The system provides the user with 3 tries for each task and includes a timing mechanism. In order to confirm the response, the user is required to select the proper answer based on what they believe to be correct.

### 4.5.6 Implementation of Avatars and Record Speech and Description Text

The following tools have been utilised to enhance the components incorporated later on in the experimental interface:

1. Microsoft Office PowerPoint 2007, used to create the textual information connected to the communicated information.

2. Camtasia Studio by TechSmith Corporation [2] used mostly for recording female and male speech sounds where the PowerPoint presentation of the learning material is running, and then producing the video file in a AVI format (Audio Visual Interleave). The output file was a visual presentation of the learning content accompanied by the spoken information by the Avatar.

The presentation process was timed consuming mainly because occasionally the number of frames became larger and the machine became suspended. To resolve this problem, each information was divided into 2 or 3 parts each of which was separately processed via improvement tools such as Crazy Talk [3] which was utilised in the production of the avatar’s torso (head and shoulders), due to the requirements of this experimental condition entailing the necessity for movements of the head and eyes symbolising human expressions. Crazy Talk provides better 3D facial orientation, face profile for all kinds of creatures, enhanced hair mesh for natural head movement. Additionally, Crazy Talk transforms the action of photos and images into talking animated characters and brings them to life with real-time actions; turning any PC into a face animation movie studio. Within Crazy Talk, one can also apply many types of special effects to enrich one’s animation. Figure 1 and Figure 2 show the audio-visual message used in the VI and RI experimental system. It can be seen that the message format differed from that used in the previous experimental condition. This experiment assumed that communicating this type of information and knowledge simultaneously has the potential to enhance the usefulness of e-assessment interface. Thus, the order, in which information and knowledge was communicated to the user, differed considerably in the implementation of the AI condition.

### 4.5.7 Pre-Experimental Questions

Thirty users participated in the study, mostly students of undergraduate in higher education. The questions were related to the subject matter to be examined as the e-assessment content. Everything is written clearly and is easy to understand. The questionnaire and questions about the time limitation for the users, task order and to answers to specific questions were layout in order to reduce the possibility of users making mistakes. The experiment was clearly explained to all users before it started via the pre-experimental information for users, which included profiling and to get their viewpoints in regard to the use of images with avatars, recorded speech, images and description text in e-assessment interfaces. In this questionnaire, users were asked to:

1. Provide personal and educational information.

2. Record previous knowledge about each of the metaphors and e-assessment.

3. Provide their opinions about the use of multimodality in e-assessment interfaces.

## 4.6 Variables

The types of variables considered in this experiment were: independent, dependent and controlled variables. The tasks were similar for all users. The level of difficulty of the subject matter was varied according to moderate, difficult and very difficult. The time distribution was not the same for all tasks and depended on the user, the users were aware of the tasks that would be provided to them.

### 4.6.1 Dependent Variables

Dependent variables were observed:

DV 1: Question answering time: measured via the time spent by users in answering the required questions.

DV 2: Number of correct answers: measured by the percentage of correct answers attained through users in answer to the required questions.

DV 3: User satisfaction: measured through user ratings in reply to the satisfaction questionnaire

DV 4: Preferred e-assessment interface: attained via calculating which user interface users preferred.

| **Dependent Variables code** | **Variables** | **Measure** |
|---|---|---|
| **DV 1** | Question answering time | Efficiency |
| **DV 2** | Number of correct answers | Effectiveness and users learning performance |
| **DV 3** | User Satisfactions | Satisfactions |

**Table 4.2: Dependent variables**

| **Independent Variables code** | **Variables** | **Condition 1** | **Condition 2** | **Condition 3** |
|---|---|---|---|---|
| **IV 1** | Presentation mode | AI | RI | DI |
| **IV 2** | Question complexity | Easy | Moderate | Difficult |

**Table 4.3: Independent variables**

### 4.6.2 Independent Variables

Independent variables are the factors in the experiment that are the cause of outcomes. These variables include the following:

IV 1: Presentation mode: the experimental e-assessment condition offered three different modes for the presentation of the learning material; Avatar with images (AI) description text with images (DI) and Record speech with images (RI).

IV 2: Question complexity: the experiments investigate usability in relation to three levels of complexity, namely; easy, moderate and difficult.

## 4.7 Questionnaire

Eight pages were contained within this questionnaire; with the first element detailing the instructions to the user and providing a space for the input of personal details (see Appendix B1). At this stage, users start by reading the instructions which are similar for each of the three conditions. Instructions for completing the tasks are similar for all interfaces. Included with the standard satisfaction statements are additional statements provided with each interface in accordance with the type of modalities prevalent within the respective interface. The final page of the questionnaire was divided into two sections: the first requiring users to select the condition they found most enjoyable, with a provision for them to rank them according to their experience; and the second section allowing for users to detail any problems that they experienced along with any other comments or suggestions.

### 4.7.1 Sampling

For equal distribution of the assessment among the three interfaces, the number of users should be a multiple of 15. Thus, the test sample consisted of 30 users who were both undergraduate and postgraduate students at De Montfort University. All of them were volunteers and used the multimodal experimental interface for the first time because the involvement of expert users in this subject would impact the experimental outcome because most probably they would rely on their previous knowledge in answering the questions and thus the impact of the tested experimental e-assessment interface on the users' performance will be negated. It also gives users the ability to provide observations and feedback about the questions. The users participating in the multimodal experiment had to be fluent English speakers because the correct pronunciation was required for using this experiment to reduce the occurance of speech recognition errors, which could also affect the results negatively.

### 4.7.2 Data Gathering

The prime focus of the data gathering exercise was predominantly the questionnaires and the experimental observations. Two types of data were products of the study measurements, namely: objective data and subjective data. The former type represents the outcome from the system used for observation which was also used to measure efficiency and effectiveness, whereas the latter type represents an outcome from the questionnaire which was used in the evaluation of user satisfaction. To guarantee that the results would be accurate, data was collected automatically by the system due to the installation of a built-in system capable of such feats. This demonstrated greater precision in the measurement of efficiency. This applied to the data associated with effectiveness, where the accuracy of users' answers was gained and the number of successfully answered questions for every user was counted via the system. The students who participated in the study were typically from a higher education background and were of diverse ages, backgrounds and gender.

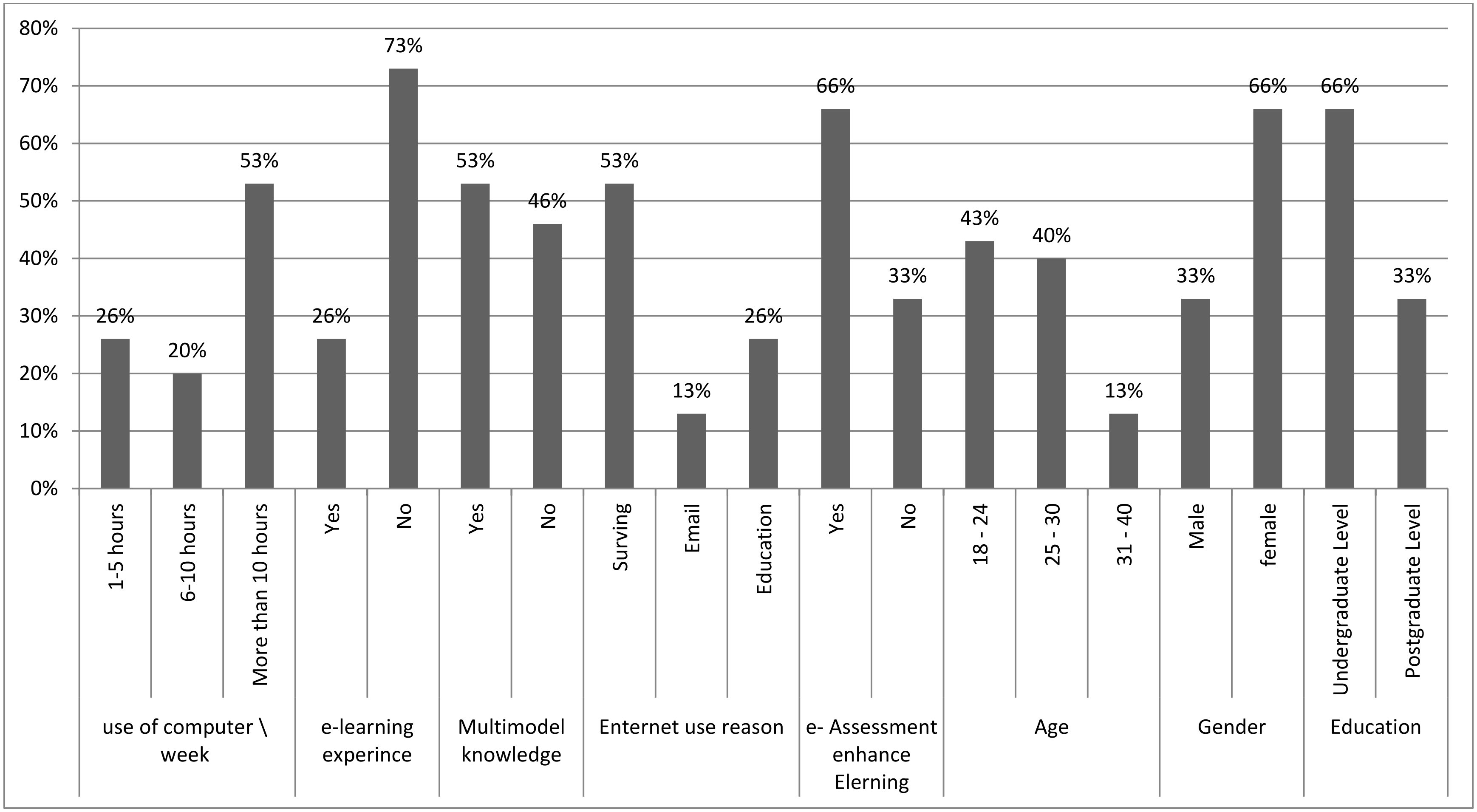



**Figure 4.2: Users' profile in terms of age, gender, educational and prior experience of users**

## 4.8 Users Profiles

This experimental stage involved 30 users; who contributed individually to the experiment. Users also gave their prior experience and views to the pre-experimental questionnaire which were analyzed to determine their personal and educational information related to the use of multimodality in e-assessment. Figure 4.4 shows that 43% of the students participants had an average age of between 18 - 24, and another 40 % of between 25 – 30 years, and 13 % between 31 – 40 years. There were 66 % female and 33% male users. As for the educational level, 66% were undergraduates and 33 % postgraduates. As can be seen in Figure 4.2 53 % of the users were using computers on a weekly basis for longer than 10 hours, 26 % for six to ten hours and 20 % for one to five hours. With respect to e-Learning experience 26 % of the users were experienced in e-Learning methods, frequently used e-learning websites or software on a weekly basis, however, for different periods of time, and 73 % of the users had no knowledge about e-Learning systems. Regarding Multimodal knowledge 53% of the users had no knowledge at all, and 46% had good knowledge about Avatars and recorded speech. Concerning internet use 53 % use the internet for surfing, 46% for education and 13% for email. With respect to e-assessment enhancing e-learning 66 % agreed with this idea and 33% disagreed. The frequency of user profiles, raw data can be found in the Appendix (B1).

## 4.9 Results and Analysis

In total, 30 users took part in the study, which took over 4 weeks, mainly in the De Montfort University library. Both assessments were distributed randomly, but were similar for each volunteer. The obtained outcomes were analysed in terms of the use of multimodality use in the experimental interface, correctly answered questions (effectiveness and users‘ learning performance), answering time (efficiency) also user satisfaction. The nonparametric Chi-square test was used to examine the significance of differences in terms of categorical data such as users’ views [124] . In addition, Wilcoxon Signed-rank test was used as the non-parametric equivalent of the dependent t-test [121] to execute follow-up pair-wise comparisons across the experimental condition in this experiment. The significance level used in these statistical analysis tests was α = 0.05 which shows the existence of dissimilarity if the p-value was less than that value.

### 4.9.1 Efficiency

The timing notes were collected from the users on the condition, via taking the mean different time users spent in each interface. The efficiency of each experimental interface was measured using the time taken via users to answer questions connected to the e-assessment offered by that interface. This measure was considered for every question and according to question type (moderate, difficult, complex) and different multimodal condition. In condition RI the mean time was 69.6 seconds, for Condition AI 47.7 seconds, and condition DI was 79.8 seconds, the time spent in condition DI was clearly higher than in the other condition as depicted in Figure 4.6.

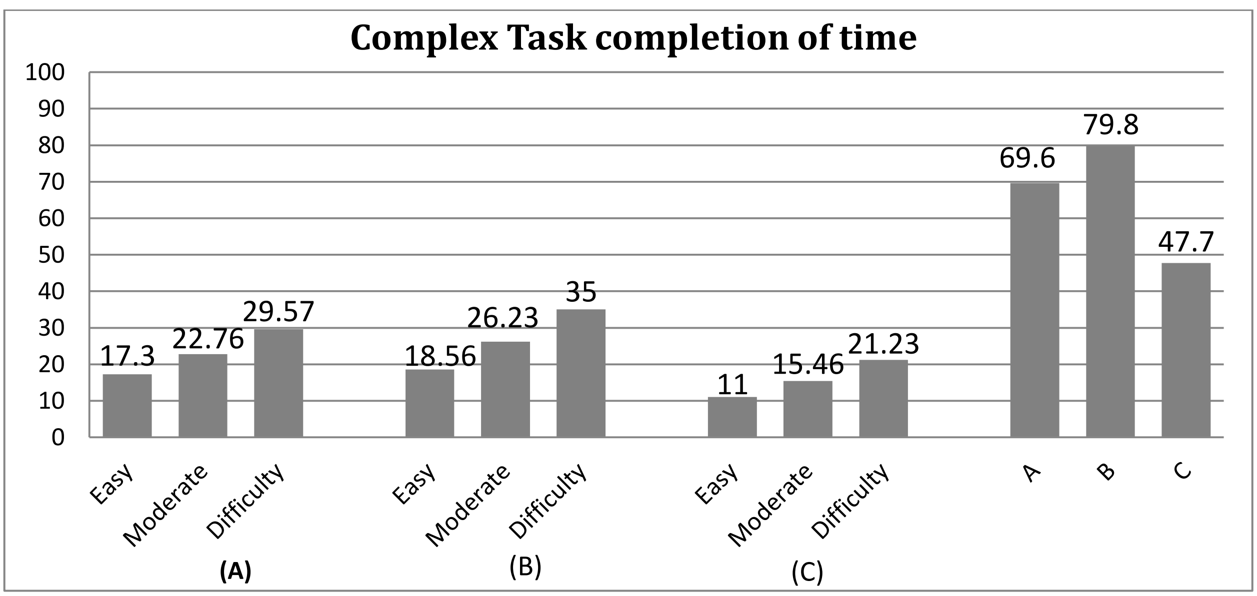



**Figure 4.3: Mean values of time taken by users to answer all questions in experimental by each condition, (A) record speech with images, (B) desecration text with images, (C) avatar with images**

### 4.9.2 Task Complexity

Figure 4.3 shows the mean values for task accomplishment, according to the three task complexity levels, using the AI, DI and RI experimental systems. These questions were designed to increase in difficulty and were divided into three easy, three moderate and three difficult of each interface. It can be noted that the AI condition outperformed the DI also RI condition in all levels of complexity for task completion time of the difficult questions it was slightly lower 47.7 minutes to answer all questions. In particular, the use of multimodal metaphors has been shown to have a considerable impact on task completion time and the average of task accomplishment for each interface. In easy questions, the mean value for AI condition time (11 minutes) was slightly lower than those of the DI and RI condition. While the mean values in moderate and complex task for AI condition were 15.46 and 21.23 minutes, considerably lower than that for DI and RI. In fact, the use of multimodal metaphors enhanced the user‘s efficiency to perform complex tasks, because the analysis of the users' performance showed that the mean task accomplishment per unit of time was less in AI condition than in the DI and RI condition, the raw data of the average time spent can be found in the Appendix (B5).

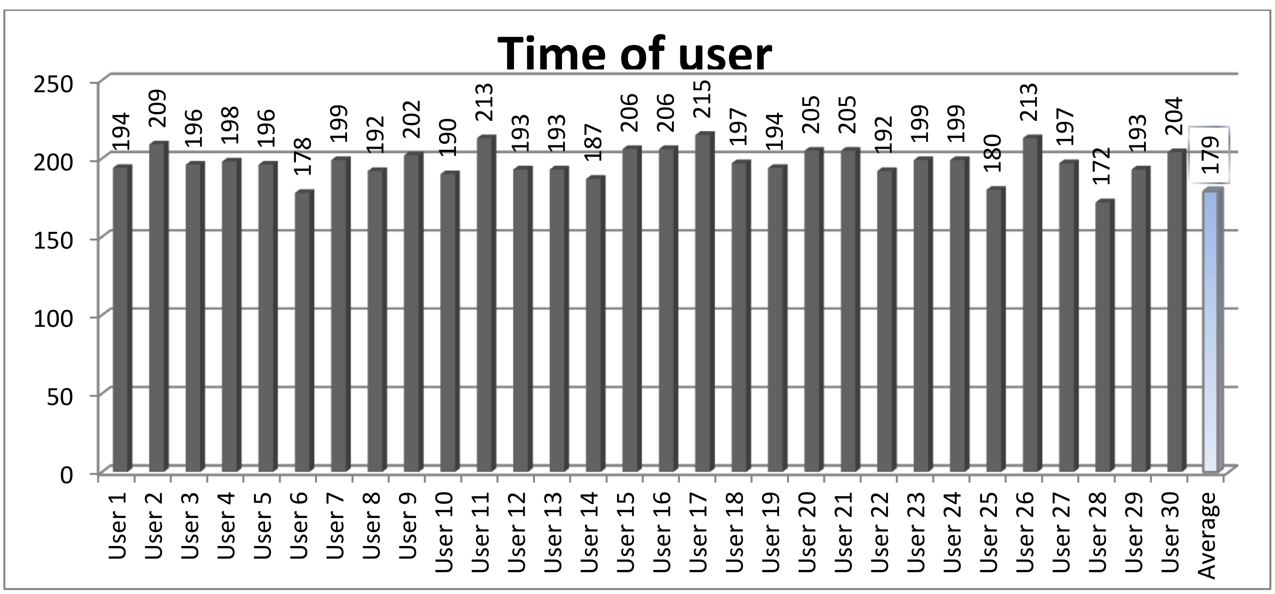



**Figure 4.4: Mean answering time for questions related to each question the experimental e-assessment condition**

### 4.9.3 Individual User

Figure 4.4 shows the entire time taken for each user to answer the required task for the experimental condition. In each task, it can be noted that users were quicker in providing the answers when Avatar with images and slower for the other multimodal combinations. In all, the average time for all users was 179 seconds, however, the average answering time for all tasks was found to be medium. The quickest time was 172 seconds by user 28 who was the quickest of the users and the slowest was 215 seconds by user 18. As far as the experimental interfaces are concerned, users were more efficient using the AI interface averaging (477 seconds) seconds per user as opposed to the use of RI (696 seconds) and DI (798 seconds) (see Figure 4.4) The raw data of the spent time can be found in the Appendix (B5).

### 4.9.4 Effectiveness

The number of correct answers was used to determine the effectiveness of the experimental interface. This measure was a total for all the questions. As for question complexity (easy, moderate and difficult) figure 4.6 shows the score of correct answers to all questions and for each user. It can be seen that users' performance was the highest with the implementation of the Avatar with images. Figure 4.5 shows the percentage of correct answers to all questions in relation to question complexity. The raw data for the correctness of users' answers can be found in the Appendix (B6).

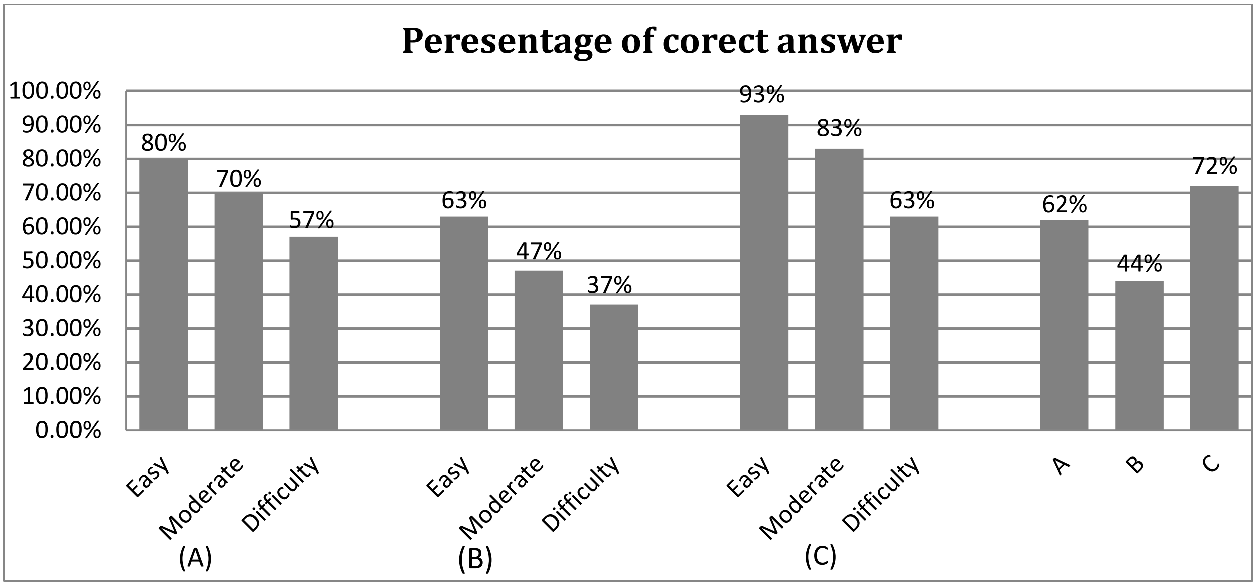



**Figure 4.5: Percentage of correct answers achieved by users question complexity in experimental by each condition, (A) record speech with images, (B) description text with images, (C) avatar with images**

The results reveal that the users of the AI outperformed DI and AI users in terms of correctness of answers to all questions as well as to each complexity level. In easy questions, the mean value of AI condition for correctness of answers at 93% was clearly higher than those of the DI and RI condition. While the moderators' questions were 83% also clearly higher than those of the DI and RI and for complex questions about the AI condition was at 63%, again considerably higher than that for DI and RI. In fact, the use of Avatar with images enhanced user efficiency to perform complex tasks, because the analysis of the users' performance showed that the mean task accomplishment per unit of correctness of answers was higher in AI condition than the DI and RI condition.  The raw data for the correctness of users' answers can be found in Appendix B6.

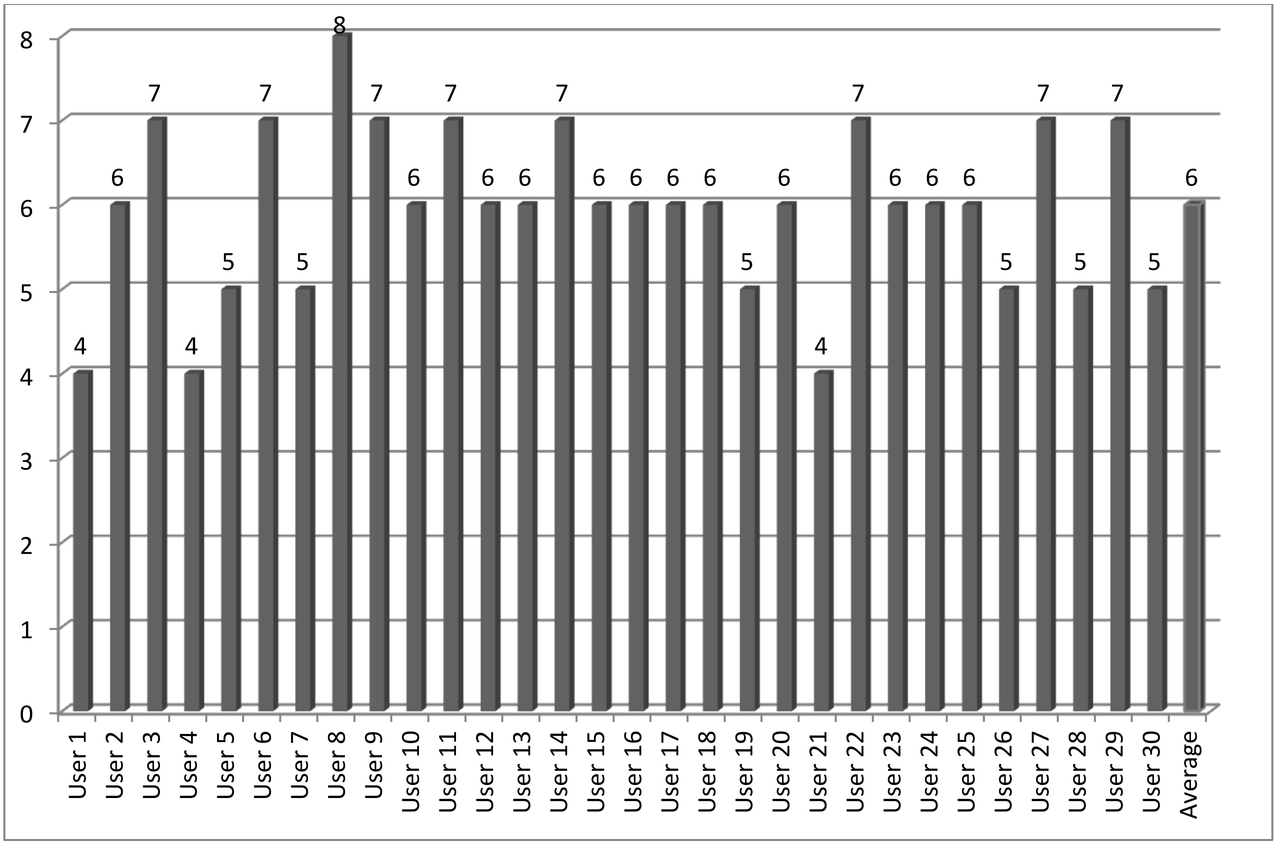



**Figure 4.6: Number of correct answers achieved by users for all questions**

### 4.9.5 Individual User

Figure 4.6 shows the total number of correct answers achieved by each user in the experimental groups. None of these experimental users were able to reach a similar performance level of an average of 8 correct answers. It is worthy to note the average was 6, the questions of the most correctly answered by users (3, 6, 9, 11, 14, 22, 27 and 29) achieve 7 correct answers. On the other hand, the weakest performers in the experiment (Users 1, 4, 21) scored only 4 correct answers; this may be due to the fact that the users don't have knowledge of multimodality. Using multimodal metaphors like avatar with images result in capturing the users' attention and communicating the learning material assisted users to answer questions correctly. Thus, such a combination could assist in enhancing learners' performance in responding to different assessment questions. More details about the correctness of users' answers to the learning evaluation questions can be found in the Appendix (B6).

### 4.9.6 Users Satisfaction

Users' responses to the SUS questionnaire (10 statements) measure the user's enjoyment and pleasure and assist with understanding the impact of the multimodal metaphors in enhancing users' satisfaction after they have had the chance to use all three experimental interfaces. The SUS questionnaire used a Likert five-point scale [125]. SUS scoring process to calculate the satisfaction score for every user in each condition and also to determine which interface; users enjoyed the most and were motivated to continue using. The five-point Likert scales were structured with 1 representing "strongly disagree" and 5 representing "strongly agree". The occurrence of the user agreement for every statement was accumulated to discover the users' judgment towards the interface, multimodal aspects and learning experience. The users on the AI condition thought the learning was better than the DI and RI condition. It was easier for them to identify the correct answer communicated by an avatar with the image. The raw data for user satisfaction can be found in the Appendix (B7, B8, B9).

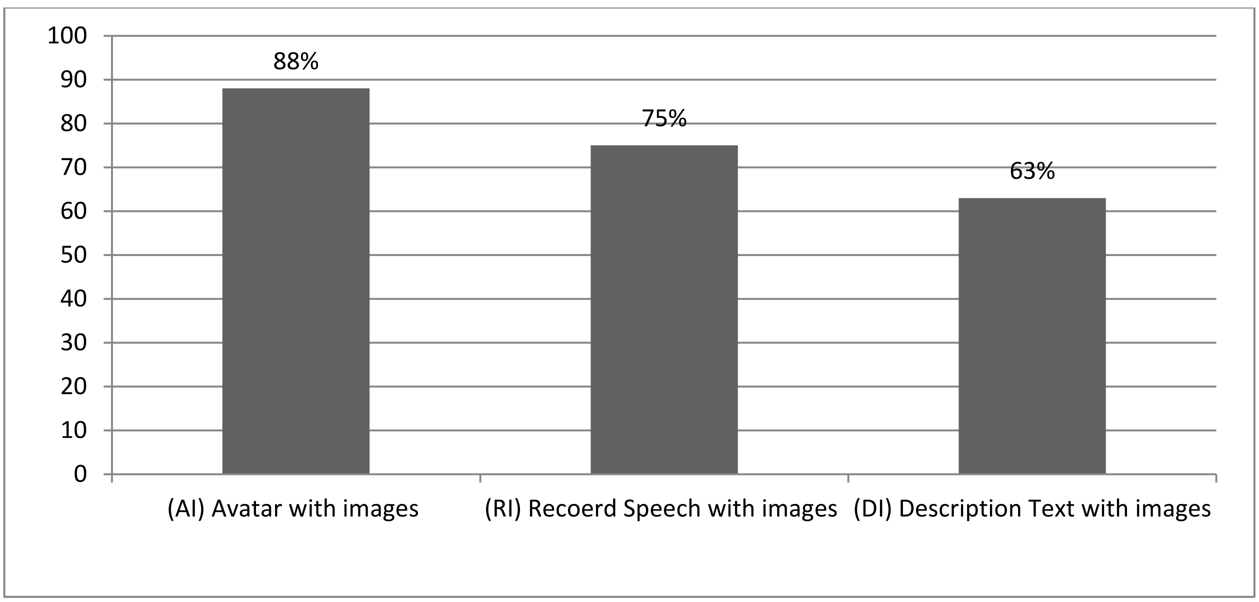



**Figure 4.7: SUS Scale result**

## 4.10 Normality Test

In statistics, statistical methods are based on various underlying assumptions, normality tests are used to determine if a data set that has a random variable is normally distributed. In several statistical analyses, normality is usually without any empirical evidence or test. However, normality is significant in several statistical systems. When this supposition is violated, interpretation and inference might not be reliable.

Consequently, before conducting the ANOVA test, normality was tested using SPSS software to analyse the time and correctness of the data in this experiment, the tables 4.3, 4.4, 4.6 and 4.5 concluded the test outcomes. In Kolmogorov-Smirnova distribution the RI was, D (189) = 0.235, $P < .05$, and D (201) = 0.178, $P < 0.05$, and D (202) = 0.171, $P < 0.05$, this appears to be normal, the DI, D (149) = > 0.522, $P < 0.05$,and D (175) = 0.314, $P < 0.05$ and D (213) = 0.132, $P < 0.05$, appears normally distributed, and in AI (205) = 0.162, $P < 0.05$, and D (139) = 0.606, $P < 0.05$ and D (120) = 0.782, $P < 0.05$ also appears normal. Thus, the result as stated here is that all variables are significantly normally distributed.

Chi-square tests showed a significant difference between AI and DI with regard to the rate of task completion ($p = 4.55 > 0.05$). In fact, there was a difference in task completion between the two interfaces. Thus, it can be said that the presence of difference has a positive effect on the contribution of multimodal metaphors (Avatar and images). The task completion rate for AI was lower than that for RI ($p = 0.0 < 0.05$). In addition, the task completion rate for DI was lower than that for RI ($p = 0.0 < 0.05$)

**One-Sample Kolmogorov-Smirnov Test**

| | **(RI)** | | | **(DI)** | | | **(AI)** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **T1** | **T2** | **T3** | **T4** | **T5** | **T6** | **T 7** | **T8** | **T9** |
| N | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 |
| Normal Mean | 17.30 | 22.76 | 29.56 | 18.56 | 26.23 | 35.03 | 11.06 | 15.46 | 21.23 |
| Std. | 3.12 | 4.42 | 4.28 | 2.09 | 3.70 | 4.86 | 2.40 | 2.55 | 2.22 |

| Deviation | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Extreme Differences | .189 | .201 | .202 | .149 | .175 | .213 | .205 | .139 | .120 |
| Positive | .118 | .201 | .202 | .090 | .175 | .196 | .205 | .139 | .110 |
| Negative | -.189 | -.141 | -.175 | -.149 | -.146 | -.213 | -.095 | -.139 | -.120 |
| Kolmogorov-Smirnov Z | 1.03 | 1.10 | 1.10 | .814 | .961 | 1.16 | 1.12 | .762 | .657 |
| Asymp. Sig. (2-tailed) | .235 | .178 | .171 | .522 | .314 | .132 | .162 | .606 | .782 |

a. Test distribution is Normal.

b. Calculated from data.

**Table 4.4: Test of normality for time**

In addition, a normality was conducted on the number of correct answers, where the data as summarised in Table 4.5 below. The Kolmogorov-Smirnova distribution showed the RI was D (488) = 0, P< 0.05, and D (440) = 0, P< 0.05and D (372) = 001, P< 0.05 and appeared to be non-normal. For the DI it was D (406) = 000, P< 0.05, and D (354) = 001, P< 0.05 and D (372) = 0, P< 0.05 which appeared to be a non-normal distribution and in AI it was D (537) = 0, P< 0.05, and D (503) = 0, P< 0.05 and D (406) = 0, P< 0.05, again a non-normal distribution. Thus, the overall finding here is that all variables are significantly non-normally distributed. There are no variables that are normally distributed and a non-parametric test (Friedman test) should be conducted to compare and find out the differences between the conditions.

**One-Sample Kolmogorov-Smirnov Test**

| | (RI) | | | (DI) | | | (AI) | | |
|---|---|---|---|---|---|---|---|---|---|
| | **T1** | **T2** | **T3** | **T4** | **T5** | **T6** | **T 7** | **T8** | **T9** |
| N | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 |
| Normal Mean | .800 | .700 | .566 | .633 | .466 | .366 | .933 | .833 | .633 |
| Std. Deviation | .406 | .466 | .504 | .490 | .507 | .490 | .253 | .379 | .490 |
| Extreme Differences | .488 | .440 | .372 | .406 | .354 | .406 | .537 | .503 | .406 |
| Positive | .312 | .260 | .303 | .269 | .354 | .406 | .396 | .330 | .269 |
| Negative | -.488 | -.440 | -.372 | -.406 | -.320 | -.269 | -.537 | -.503 | -.406 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Kolmogorov-Smirnov Z | 2.676 | 2.411 | 2.036 | 2.224 | 1.941 | 2.224 | 2.941 | 2.756 | 2.224 |
| Sig. (2-tailed) | .000 | .000 | .001 | .000 | .001 | .000 | .000 | .000 | .000 |

a. Test distribution is Normal.

b. Calculated from data.

**Table 4.5: Test of normality for correct answer**

### 4. 10. 1 Friedman's ANOVA Test

Friedman's ANOVA test is a non-parametric statistical test, it is used to detect and test differences between conditions when there are more than two conditions, and the same users have engaged in all condition, which is same as the experiment presented here. Also, this test can be used to test the dissimilarities among experimental conditions within subject design when the assumption of normal distribution of the data is violated [126, 127]. In addition, the Friedman test focuses on the fact that the samples should be randomly taken and independent of each other. Consequently, this test has been conducted for the time in all conditions using the SPSS statistical software and the output of the result is shown in the Table 4.6.

**Ranks**

| | **Mean Rank** |
|---|---|
| AI | 2.50 |
| RI | 2.00 |
| DI | 1.50 |

**Test Statistics**

| | |
|---|---|
| N | 30 |
| Chi-Square | 19.780 |
| df | 2 |
| Sig. | .000 |

**Table 4.6: Friedman's ANOVA test results for correct answer**

Due to the fact that the p-value = 0.00 ≤ 0.01 = α, the null hypothesis is rejected, and we can say that at the α = 0.01 level of significance, there is enough proof to elicit that a difference exists in the true mean correct answers recorded for the three experimental conditions.

### 4.10. 2 Wilcoxon Signed-rank Test

Wilcoxon signed-rank test focuses on the differences between scores in the two experimental conditions to be compared with this experiment. Consequently, each experimental condition has been tested separately, since three conditions are examined here, once these differences have been calculated they are ranked and the result is assigned to each rank, thus the comparison distribution for the interface is achieved through AI comparison with interface RI and AI with interface DI and DI with interface RI. The outcomes of the test are shown in Table 4.6. The first table shows the positive and negative ranks. While the positive rank indicates the users score compared to the other condition, the negative rank shows the opposite. The first result of a comparison between recorded speech with images (RI) and Avatar with images (AI) was 6 out of a potential 30 positive rank, whereas in (AI) compared with (DI), there were 22 positive rankings out of 30 rankings, thus the score was greater in (AI) than for (DI). In the comparison between (RI) and (DI) 14 of the 30 users gave a positive rank to (RI). There also ties in the ranking in the table, which means there are some users that gave the same score for each condition. The table shows as well the mean number of negative and positive ranks as well as their sums. Below the table refers to the comparison outcomes represented arithmetically, the significance stage used in these statistical tests was $\alpha = 0.05$ indicating the existence of significant dissimilarity if p-value was less than that value [126]. The Table 4.6 below contains the output of the test statistics of correct answers. The value of Z was 1.65 and this value is significant at p= 0.98 because this value is focused on the positive rank; therefore, the correctness increased in the (AI) condition more than the (RI) ($Z= -1.65$, $p< 0.05$). In the case of the comparison of (DI) with (AI), Z was -3.621 and this value is significant at p= 0.0, and because this score was related to negative ranks, the users provided more correct answers in the (AI) condition than they did in (DI) ($Z = -3.621$, $p<.05$). The value of Z in the comparison of (RI) with (DI) was -2.851 and this value is significant at p=.004. Because this value is related to the negative rank, consequently, the correctness were significantly lower in the (RI) than in the (DI) ($Z= -2.851$, $p< 0.05$).

**Rank**

| | | N | Mean Rank | Sum of Ranks |
|---|---|---|---|---|
| **RI – AI** | **Negative Ranks** | 16a | 10.88 | 174.00 |
| | **Positive Ranks** | 6b | 13.17 | 79.00 |
| | **Ties** | 8c | | |
| | **Total** | 30 | | |
| **AI – DI** | **Negative Ranks** | 2a | 13.75 | 27.50 |
| | **Positive Ranks** | 22b | 12.39 | 272.50 |
| | **Ties** | 6c | | |
| | **Total** | 30 | | |
| **RI – DI** | **Negative Ranks** | 4a | 5.50 | 22.00 |
| | **Positive Ranks** | 14b | 10.64 | 149.00 |
| | **Ties** | 12c | | |
| | **Total** | 30 | | |

a. RI < AI     b. RI = AI     c. AI > DI     d. RI < DI     e. RI = DI

f. RI > AI     g. AI < DI     H. AI = DI     i. RI > DI

| | RI - AI | AI - DI | RI - DI |
|---|---|---|---|
| **Z** | -1.652b | -3.621b | -2.851b |
| **Sig. (2-tailed)** | .098 | .000 | .004 |

a. Based on positive ranks     b. Based on negative ranks     c. Based on negative ranks

**Table 4.7: Wilcoxon signed-rank test results for correct answers**

## 4. 11 Discussion

The aim of this chapter was to investigate how a combination of metaphors assists users while engaged in multimodal e-assessment systems. This applies to three quantitative aspects: time spent by users;

accuracy; and user satisfaction levels. Through the experiments, it was noticed that the users who participated in the valuation of the e-assessment in the AI interface were confident in performing most of the tasks, because they have had prior experience with avatars.

The results indicate that the AI (Avatar with images) condition was more successful than the other condition in increasing usability. This was evidenced by the fact that the users made fewer mistakes while using the AI condition; their performance was increased and there was a high level of satisfaction and enjoyment. These positive results were regardless of the time taken for the tasks. It was found in relation to H3 there were significant differences in accuracy for the different multimodal interfaces, specifically, 88% for the AI condition, and 75% for the RI condition and for the DI condition, 63% (see Figure 4.7). Accordingly, the results advocated H3.

The experiment's dependent variable measurements were suitably managed, to provide accurate results. These measurements were achieved through incorporating a fitted integral time design into all interfaces. In this way, the users in the AI achieved greater interaction with this interface. Generally speaking, the results reported were positive. Moreover, the overall test outcomes highlighted significant areas and implications for more studies, with the data analysis outcomes affording additional opportunities for research.

The outcomes from Figures 4.3 showed that the efficiency of AI, in terms of time was different in the three interfaces. This difference, mostly, caused by the existence of differences between the three interfaces in task accomplishment as shown in Figure 4.3. The existences of difference in efficiency of completing tasks were noticeable between AI, DI, and RI as was expected in H2. It is noteworthy that the time spent on DI was longer than on the other interfaces; a result supporting H2. A possible explanation for the differences in mean times is that the AI condition was designed to be more time consuming, because of the incorporation of multimedia elements that made the AI interactive with sound and video, images.

Furthermore, it is clear that the percentage of correct answers decreased, gradually adding strength to H6, from the AI condition to the RI condition and finally, to the DI condition. The majority of users had limited previous knowledge of multimodality. It should be remembered that the impact of an avatar on AI condition may explain this success. The results gave additional evidence of the impact of multimodal metaphors, represented by video, sound and speech on the AI interface, in further enhancing usability.

An overall improvement was noticeable in the performance of users on the AI condition. However, although the results in the condition RI show a degree of improvement, it was not as significant as in AI, whereas, in condition DI, the results showed a lower impact on usability. This also emphasises the power of the avatar as a tool for conveying information to users and enhancing the interaction between students and e-learning systems; a finding borne out by previous research [128]. In addition, the avatar is not a video clip; it is created on interactive elements, and communicates with the user as a reaction to what the user requests. Recently, the avatar became commonly used in several areas to improve interactivity, learning engagement and cultural issues [128].

A comparison of the scores between all condition systems rated the interface with the AI condition higher than the other conditions;, therefore, adding strength to H3. The mean score calculated for the condition RI was 62 %, compared to 72 % of AI, and 44 % for DI. Thus, the design of interface, AI was found to be the most effective for assisting users to find the correct answer. User satisfaction levels were a little higher for the AI interface than for RI interface. Thus, multimodality provided greater user satisfaction and also encouraged the users to persevere with the tasks. This enabled the users to acquire the information, where they actually enjoyed the avatar, despite the time spent on the task. H4 is therefore supported by these results.

In addition to the evidence that the AI condition was the users' condition of preference, a number of serial tests (ANOVA) were also conducted using the data, in order to establish the differences among the experimental interfaces, in terms of time taken and levels of accuracy achieved. These tests resulted in significant variations, as expected, and the hypotheses were rejected. A Wilcoxon test result also differentiated between the interface as stated above and error rates were significantly lower in the AI condition than they were the RI and DI condition.

In terms of user preference, the AI condition was the main condition of preference, followed by the RI condition, with the DI condition being the least favoured. Moreover, a user's order of preference for the different condition further supported the AI interface, which achieved the highest mean user satisfaction, at 88% in comparison to 75% for the RI condition and 63% for the DI condition. This chapter has also investigated three various modes of multimodality, the data analysis proved, unequivocally, that AI condition was, ultimately, the superior interface, compared to the other two conditions. The obtained results were used to compare these experimental conditions in terms of efficiency, effectiveness and user satisfaction.

## 4. 12 Summary

This chapter describes a comparative valuation study carried out to assess the impact of incorporating avatars with images, description text with images and recorded speech with images into e-assessment condition and not only to evaluate the usability of e-assessment, but also to investigate assisting users with different levels of ability. In order to the aims of experimenting with combinations of multimodal metaphors, three experimental interfaces were implemented and compared. The assessed usability measures included efficiency (in terms of task completion quickly), effectiveness (in terms of tasks completed correctly answer) and user satisfaction. The obtained outcomes demonstrated that facial expression avatars were considered as a positive contribution in the e-assessment interface. These findings suggest that the adoption of avatars played an enjoyable and attractive role in delivering information to users. Moreover, the general test outcomes highlighted several significant areas and implications for future research. In addition, the results established that the AI condition proved more efficient, more effective and more satisfactory, as opposed to the two other conditions in the e-assessment condition investigated.

Consequently, it is essential for designers of e-assessment condition to be aware of the potential that multimodal interaction has to enhance visual communication of knowledge. Thus, the following chapter describes the third phase of this experimental programme that looks at interfaces from the user acceptance point. The third experiment also evaluates the users' acceptance under usability.

Finally, this study recommends some empirically derived guidelines for incorporating expressive avatars in e-assessment condition. The description and discussion of these guidelines are introduced in chapter 6.

# Chapter 5

# Experimental Phase III: The Role of Expressive Body Gesture and Earcons and Auditory Icons in E-assessment Interfaces

## 5.1 Introduction

This chapter leads on from the previous chapters that examined the effectiveness of multimodal metaphors from a participant's perspective. The previous experiment (Chapter 4) demonstrated how the use of Avatar with images outperformed the other multimodal combinations trialled in the last experiment in terms of enhancing users' learning performance. Moreover, the AI interface proved valuable in e-assessment and had noticeable effect in reducing error rates, in the topic sample. However, the majority of the users chose the Ai interface as the condition of preference, because the users preferred a video and sound, compared learning contents passively.

## 5.2 Aims

In line with the overall aims of the study (chapter one) this experiment aims to achieve the following:

1. To examine the impact of the use of more advanced metaphors, namely, a combination of avatars with full body gestures, earcons and auditory icons on the effectiveness of the e-assessment interface.

2. To investigate the most effective metaphors for specific types of assistance in the e-assessment process, types of assistance include, for example, involved thinking and explaining questions.

3. To investigate the usability in terms of efficiency (time taken to complete tasks) and user performance (score). A post-experiment questionnaire is conducted to derive user satisfaction. Here a usability scale (SUS) is employed.

4. To measure the recall and recognition ability of users when engaged avatars with full body gestures, earcons and auditory icons.

5. To investigate the implications of varying degrees of difficulty in the questions for each metaphor and how this has an impact on overall usability.

6. Determine the improvement on learning performance of a combination of the three metaphors.

## 5.3 Tasks

In order to achieve the above mentioned aims the following tasks were required to be accomplished:

1. Design and implement of an experimental assessment, assist condition that employs avatars in a different method to that applied in the experiment with the addition of avatars body gestures earcons and auditory icons as non-speech auditory memos to link definite structure assessment assist types.

2. Empirical evaluation of the Auditory Avatar Body Gestures interface by one group of users.

3. Measuring the effectiveness by calculating the percentage of correct answer e-assessment interface, to measure the users‟ learning performance.

4. Measure the user’s involving of testing non-speech, metaphors by users’ learnability to involve with presenting assessment, assist types, to see out which interface is more attractive for the users.

5. Measure the satisfaction of users by their answers to questionnaire dedicated to assess users' attitudes in relative to the applied e-assessment interface.

## 5.4 Hypotheses

In chapter 4 the only obvious attribute that proved to be valuable is the AI condition, but the question that this experiment seeks to establish is which multimodal attribute is responsible for improving the usability stage and the users’ learning achievements in the e-assessment interface? Based on this question the following hypotheses were derived:

H1: The earcons and auditory icons will have an impact upon the enhancement of the achievement stage of the Auditory Avatar Body Gestures interface in terms of assessment assist types and in terms of the number of tasks completed successfully; of both types of question, namely recall and recognition.

H2: Users of the Auditory Avatar Body Gestures interface will increase user involvement through assessment, assist types when information is communicated via earcons and auditory icons.

H3: Users of the experiment will express positive views towards the use of earcons and auditory icons in terms of reduced irritation and frustration, and improved usefulness and concentration.

H4: Users of the Auditory Avatar Body Gestures interface will benefit from the non-speech sounds used assessment assists types.

H5: Users will be satisfied with the Auditory Avatar Body Gestures experiment.

### 5.5 Experimental Condition

As stated earlier, this chapter aims to investigate the role of earcons, auditory icons and Avatar Body Gestures context in improving users' attention and interest as well as usability of e-assessment assist systems. The AI condition used in previous experimental work demonstrated better performance compared to other condition regarding usability and user achievement levels. The Avatar Body Gestures condition was found to be as usable as the Face to Face Condition with respect to both efficiency and effectiveness in assessment recall and recognition questions. It has been shown in these sections [96, 103, 104, 129, 130] that previous experimental studies showed the potential of multimodal metaphors in enhancing the usability of condition and users' performance, thus, integrated earcons have been proven by several researchers to be useful in enhancing the usability of systems [131]. Moreover, auditory icons as environmental sounds were successfully used to communicate information in user condition [77]. In addition, Avatar Body Gestures together with earcons and auditory icons condition provides a more realistic interaction in human computer interfaces. This assisted in making the learning process more fun to users and increased their interest, motivation, as well as improving learnability. Table 5.1 explains how earcons and auditory icons were used in the interface to capture users' attention towards the key parts of the learning content while communicating information to assist users by Avatar Body Gestures full body. This study revealed there are 6 diverse types of assist which are: Error, Comment, Involved thinking, Explain question, Suggestions and Mark. Thus, three types of multimodal interaction metaphors were incorporated in this interface: visual- (text which is assist type, content), audio-visual- (speaking avatar with body gestures) and metaphors (earcons and auditory icons).

| | | | | | | |
|---|---|---|---|---|---|---|
| | **Ability of Assistance Level** | | | | | |
| | **High** | | **Medium** | | **Low** | |
| **Assistance type** | Error | Comment | involved thinking | Explain question | More suggestions | Mark |
| **Auditory Icons** | √ | √ | √ | √ | √ | √ |
| **Earcons** | √ | | √ | | √ | |

**Table 5.1: Mapping among the important level of assessment Types and non-speech sound used in auditory avatar body gestures condition**

### 5.5.1 Assistance Types

There are six types of assessment that communicate information to users through interfaces. The first type of assistance is related to Error where the instructor determines where errors have been made (or where work is correct). The second type of assistance is Comment where the instructor represents techniques or procedures the users might not have used suitably or correctly. The third assistance type is about involved thinking which is involving users in some thinking relative to what they have written? The fourth assistance type is Explain questions, where questions or concepts the users have not exactly understood are explained. The fifth assistance type is Suggestions whereby the user is provided with suggestions for further study or reading. The final assistance type is Mark that is used to justify grades.

### 5.5.2 Implementation of Non-speech Auditory Metaphors

Earcons used in this experiment were utilised to communicate the correct answer to the question when spoken by the Body gestures avatar. The aforementioned six types of assistance were grouped in three levels in terms of their ability to help; high, medium and low. Each of these levels were represented by a rank as follows: 1 for low, 2 for medium and 3 for high. This ranking refers to the potential of each metaphor (earcons and auditory icons) to assist in communicating the correct answer. The tone used in these earcons was generated through visual music [220] and three different single-meaning earcons were designed. In a simple and meaningful format, the design of these musical stimuli was based on the guidelines for the creation of earcons [132, 133] where the sound of a drum instrument was chosen to play a dissimilar number of notes to link the required auditory messages. It can be seen that the first earcon consisted of only one note to communicate low ability, while the second earcon consisted of two increasing notes to show medium rank. Moreover, these earcons were short and helpful in the presentation of the provided oral concept to communicate assessment types were the representation of these aspects, via auditory icons, could supply natural mapping to assist the users to learn and interpret it accurately. As shown in Table 5.2, the sound of broken glass communicated that Error (type of assist) will start, and the sound of opening a bottle lid communicated that the Comment (type of assist) is about to start. Furthermore, the sound of a honking horn was used to indicate that the involving thinking (type of assist) has started, while the Explain questions (type of assist) was communicated through the sound of a closing window. The sound of an opening door indicated that the more Suggestion (type of assist) will start and the Mark (type of assist) is communicated by a clapping sound. Both earcons and auditory icons were performed in the presentation during pause intervals so that they do not interfere with the speech of the body gestures avatar.

## 5.6 Experimental Design

Usability and users' learning performance when using AI, RI, DI in the e-assessment condition was tested in the second experiment reported in Chapter 4. Thus, only one group of users will be included in this experiment to evaluate non-speech auditory sounds and body gestures avatar. Although dissimilar tasks were designed in this empirical investigation, it was believed that the obtained results could serve to explore if the earcons and auditory icons and Avatar Body Gestures are better than the metaphors used in the second experiment, meaning that this condition enhanced usability and users' learning performance. In total, 30 users participated in the experiment individually.

| **Assistance Type** | **Sound** | **Duration (seconds)** |
|---|---|---|
| **Error** | Broken glass | 18 |
| **Comment** | Opening tab of bottle | 16 |
| **Involving Thinking** | Honk | 21 |
| **Explain questions** | Closing window | 13 |
| **more Suggestions** | Opening bottle | 14 |
| **Mark** | Clapping | 20 |

**Table 5.2: Auditory icons are used in auditory avatar body gestures interface to indicate users that specific assistance type will start**

### 5.6.1 Procedure

Throughout the experiment 30 users took part in the experiment individually. The procedure that was followed in performing the experiment with each user is a follows:

1. Provide personal and educational information, relation to age, gender and educational level.

2. Record previous knowledge about assessment and multimodal metaphors.

This is followed by the presentation a short one-minute demonstration video about the tested interface. Thereafter, six assistance types provided students the opportunity to listen to the implemented non-speech sounds. The object of this training was to ensure the user's ability to understand and interpret each of these multimodal metaphors. Consequently, presentation was the same for all users (i.e. assist 1 then assist 2 then assist 3). Then, the user was instructed to perform the required tasks, the last part of the experiment was devoted to obtain user‘s opinion about body gestures and non-speech sounds as well as to provide any comments or suggestions.

### 5.6.2 Tasks

These tasks were required to be performed upon completion of assistance type and were aimed at assessing the impact on the users' achievement for a particular. Each user was asked to answer 6 questions connected to the delivered assessment type. These questions were of two types; recall and recognition. Moreover, the users were engaged in test performs to test their involvement with assistance types when using non-speech metaphors. Lastly, the final task was aimed at obtaining the users' opinion using the satisfaction questionnaire, this SUS questionnaire was composed of 10 statements of which each had a 5-point Likert scale ranging from strongly disagree to strongly agree. Additionally, users were asked to express their views towards the use of body gestures and earcons and auditory icons in terms of Irritation, Frustration, Usefulness and Concentration.

### 5.6.3 Independent Variables

1. Multimodal metaphors. In this experiment the effects of earcons and auditory icons will be investigated when incorporated with the full-body gesture avatar.
2. There are six different types of assessment. These assistance types (Error, Comments, Involving thinking, Explain question, more Suggestions and Mark) are used as independent variables.
3. Assessment Recall and Recognition question are used to evaluate the users' learning achievement attained from the information presented by the tested e-assessment interface.

### 5.6.4 Dependent Variables

1. Completion level (correct answers): This is the number of successfully achieved tasks. It is measured by the correct answers to the recall and recognition questions linked to the communicated assessment.

2. The involvement of users with the type of assessment: This was measured by the number of students who correctly indicated these features after being communicated by the non-speech sounds.

3. Users' recognition of earcons and auditory icons used: This was measured via the number of users who successfully interpreted the auditory stimuli in the context of being communicated in the experimental interface.

4. User USU: measured by users' responses to satisfaction questionnaire.

## 5.7 Data Collection and User's Profile

The same procedure which was used in the second experiment, in collecting the obtained data, including observations and questionnaires were used again in the same way, in terms of personal information and previous experience as well as their views regarding the use of body gestures and earcons and auditory icons in e-assessment interface. The tasks contributed to the evaluation of users' involvement by obtaining the data related to effectiveness and learning performance, learnability, and users' satisfaction. The test sample consisted of 30 users in the experiment who took part individually. The ages in the sample group were between 18 – 24 (13%), 25-30 (26%) and 31- 41 (60%), the gender ratio was 60% male and 40% female. With regards to educational achievement 14 users (43%) were postgraduate and 17 users (56%) were undergraduates. In reference to weekly use of computers 26% of users used a computer for between 1 and 5 hours per week, 20% for 6 to 10 hours and 53% for more than 10 hours. Of all the users 76% had excellent knowledge about multimodality. The number of users who had experience of e-Learning was 76%. The majority of users (66%) used the internet for surfing and 23% for education. Users who thought that e-assessment would enhance e-learning was only 26%. (See Appendix C-2 for frequency of user profiles).

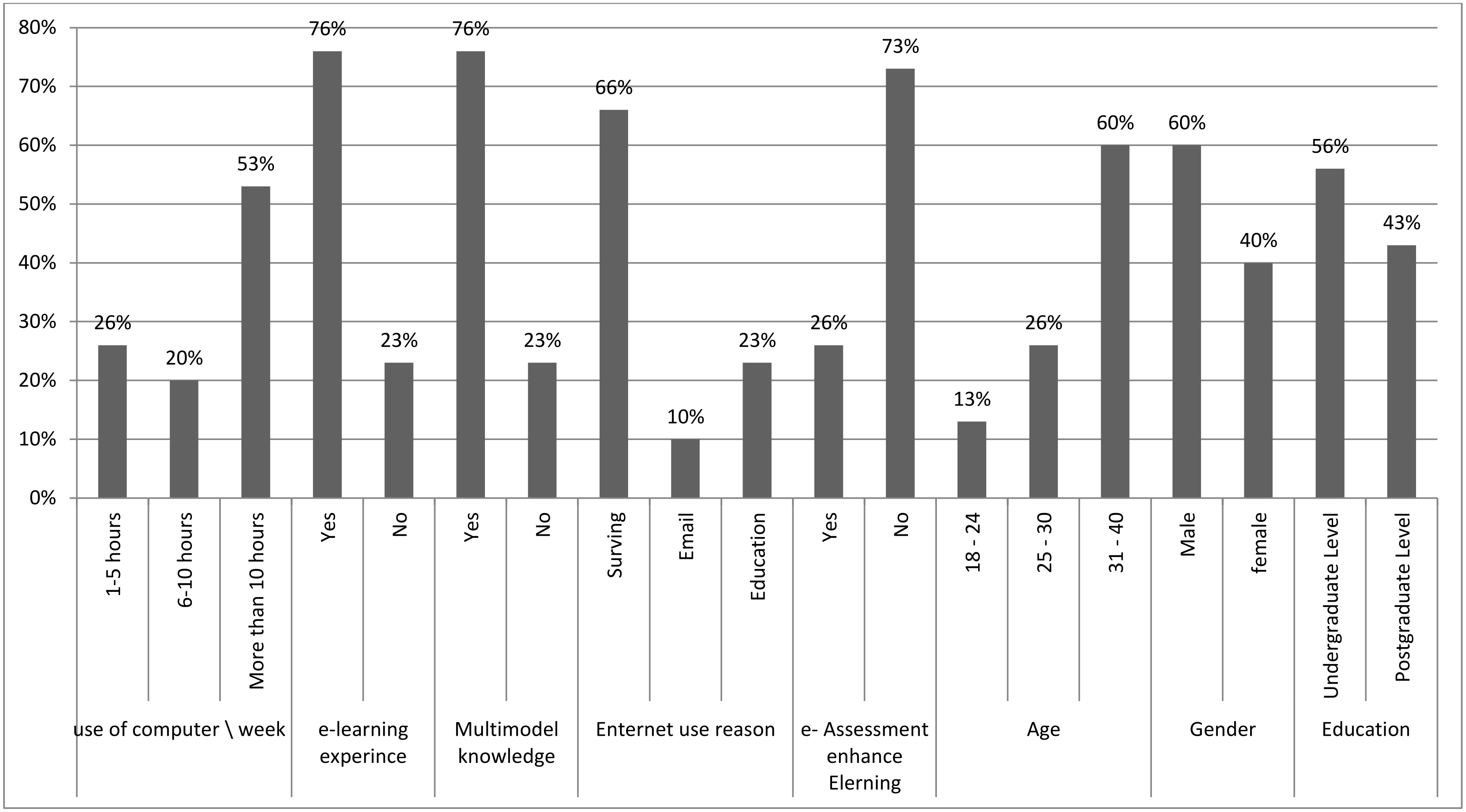



**Figure 5.1: Profile of users in terms of personal data, previous experience and Profile of users in terms of personal data**

## 5.8 Results and Analysis

The following provide descriptive and statistical analysis of the results obtained from the experiment in terms of achievement level, involving (in terms of correct and incorrect users' answers) in addition, user satisfaction, and users' views regarding the non-speech sounds that accompanied the avatar body gestures as assist. This was the results of the experimental group consisting of 30 volunteers who took part in the study. In addition, the levels of significance in students' responses were examined using the nonparametric Chi-square statistical test at $\alpha = 0.05$ indicating a significant difference when the p-value was found less than 0.05.

### 5.8.1 Achievement Stage

The numbers of correct and incorrect answers to the assessment questions were used to assess the users' achievement stage of Auditory Avatar Body Gestures. Each user answered nine questions focussed on the two assessment question types; recall and recognition. The total number of questions was 180 (30 users @ 6 questions per user) equally distributed over the two types. Figure 5.2 illustrates the percentage of correct and incorrect answers achieved by each user for the questions, grouped via assessment type and question types. Figure 5.2 illustrates that the percentage of correct answers was 78 % compared to 22% for incorrect answers. These outcomes were statistically significant (1) = 0.200, CV = 3.84, $p < 0.05$). In terms of assessment question types, Figure 5.2 explained that the percentage of successfully answered recall questions was higher than that for the recognition questions. In response to 90 questions in every type, the numbers of correct answers were (78.8%) and (87.7%) in recall and recognition questions respectively. Although users' execution was better in the recall tasks, the difference between correct and incorrect answers was significant in both assessment question types; recall (1) = 16.8, CV = 0.200 $p < .05$) and recognition ((1) = 7.4, CV = .200, $p < 0.05$).

Figure 5.2 illustrates the correct answers attained by student for every question connected to the assessment types delivered. The achievement of users was various across these questions. More specifically, the percentage of students who correctly answered questions linked to and involving thinking and error levels was 86.7% and 83.3% respectively. Nevertheless, it seems that the remaining assessment question types were difficult to answer. The percentage of correct answers for assistance types decreased to 73.7% for more suggestion, 70% for explaining question, 60% for a mark and 53.3% for comment. Table 5.3 shows that the outcomes were significant in terms of the dissimilarity among correct and incorrect answers for error, involving thinking, explain, question and more suggestion while no significance has been obtained for comment and mark assistance types.

Figure 5.3 illustrates the number of correct answers provided by each user. It can be noted that nine users (1,4,11,13,17,18,27,28 and 29) answered each question successfully while another 13 users accomplished 5 correct answers.

To sum up, it can be said that the incorporation of well-known environmental sounds and short musical stimuli along with the virtual Body Gestures was discovered to enhance and improve the delivery of the assessment content in e-assessment interfaces. Using these auditory messages can complement the role of the virtual Body Gestures which is more likely to result in capturing the users' attention. Therefore, it enhances significantly the achievement of users in successfully responding to diverse evaluation questions. More details about the data to the assessment question types can be found in Appendix C3.

| Variable | Chi-square value | Asymp. Sig. | Significance |
|---|---|---|---|
| All assessment question | .200a | .905 | No |
| **Assistance Type** | | | |
| Error | 13.333[a] | .000 | Yes |
| Comment | .133[a] | .715 | No |
| Involving Thinking | 16.133 | .000 | Yes |
| Explain Questions | 4.800a | .028 | Yes |
| More Suggestion | 6.533a | .011 | Yes |
| Mark | 1.200[a] | .273 | No |
| **Assessment questions** | | | |
| Recall | 16.800[b] | .000 | Yes |
| Recognition | 7.400[b] | .025 | Yes |

**Table 5.3: Chi-square values and significance levels relating to the achievement level**

| Variable | Chi-square value | Sig. | Significance |
|---|---|---|---|
| All Non-Speech | 15.600 | .008 | Yes |
| **Auditory Icon** | | | |
| Error | 4.800a | .028 | Yes |
| Comment | 6.533a | .011 | Yes |
| Involving Thinking | 3.333a | .068 | Yes |
| Explain Questions | 19.200a | .000 | Yes |
| More Suggestion | 10.800a | .001 | Yes |
| Mark | 16.133a | .000 | Yes |
| **Earcons for Assessment ( Ability )** | | | |
| High | 19.200[a] | .000 | Yes |
| Medium | 8.533[a] | .003 | Yes |
| Low | 3.333[a] | .068 | Yes |

**Table 5.4: Chi-square and significance levels relating to involving test 1**

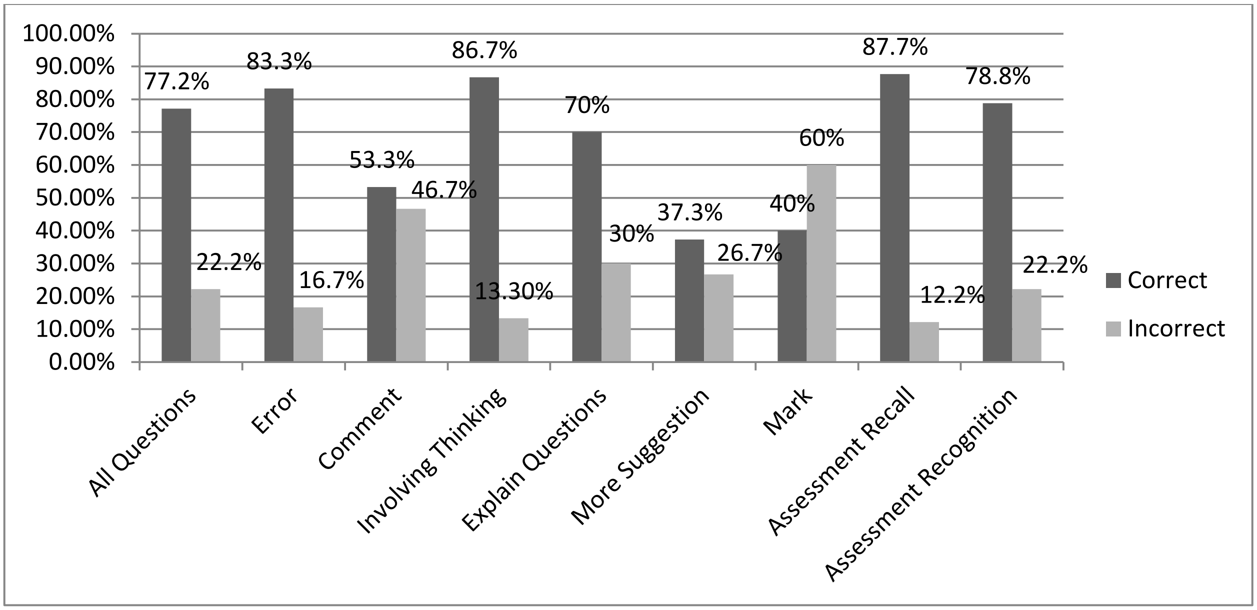



**Figure 5.2: Correct and incorrect percentages of answers achieved by users for all questions, assessment types and assessment question types**

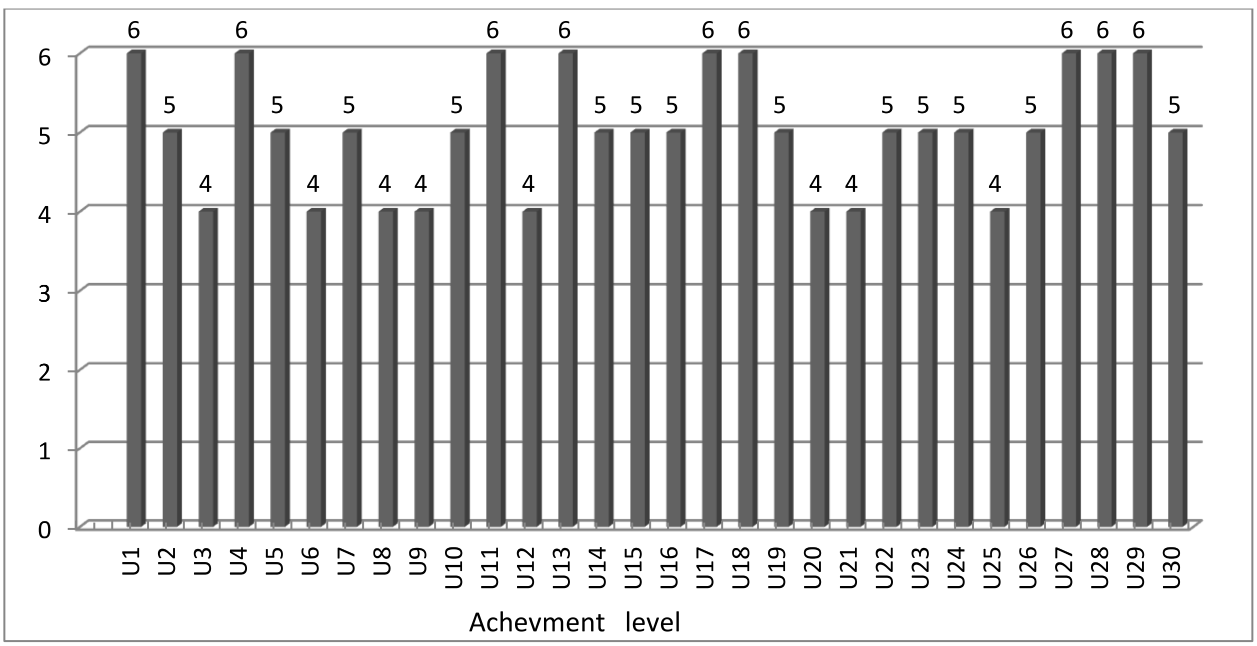



**Figure 5.3: The number of correct answers provided by each user**

### 5.8.2 Involvement

Upon completion of the achievement tasks, students were asked to do two Involving tests. Firstly, users were provided with six different assistance types and they were requested to indicate the type of non-speech sound that was the most effective. In consideration of this, the total number of questions was 180 (30 user * 6 questions per user). Figure 5.4 illustrates the correct responses of users to this task related to all non-speech sounds, earcons and auditory icons. More details are provided in Appendix C4. Figure 4 shows that the assessment that was indicated as ability as communicated by earcons and assessment types by auditory icons, was correctly recognised by students.

Therefore, statistically this result was significant ((1) = 15.6, CV = 3.84, $P < 0.05$). Also, Figure 5.4 demonstrates that the majority of users recognised correctly the assessment types communicated via auditory icons. More specifically, 93.3% (28 users) recognised Error (Type of assistance) by the sound of broken glass correctly, while 90% (27 students) accurately determined many suggestions (type of assistance) by the sound of an opening bottle and 80% (24 users) for a mark (type of assistance) by an encouraging clapping sound. This percentage decreased to 73.3% and by 66.6% for the remaining assessment types.

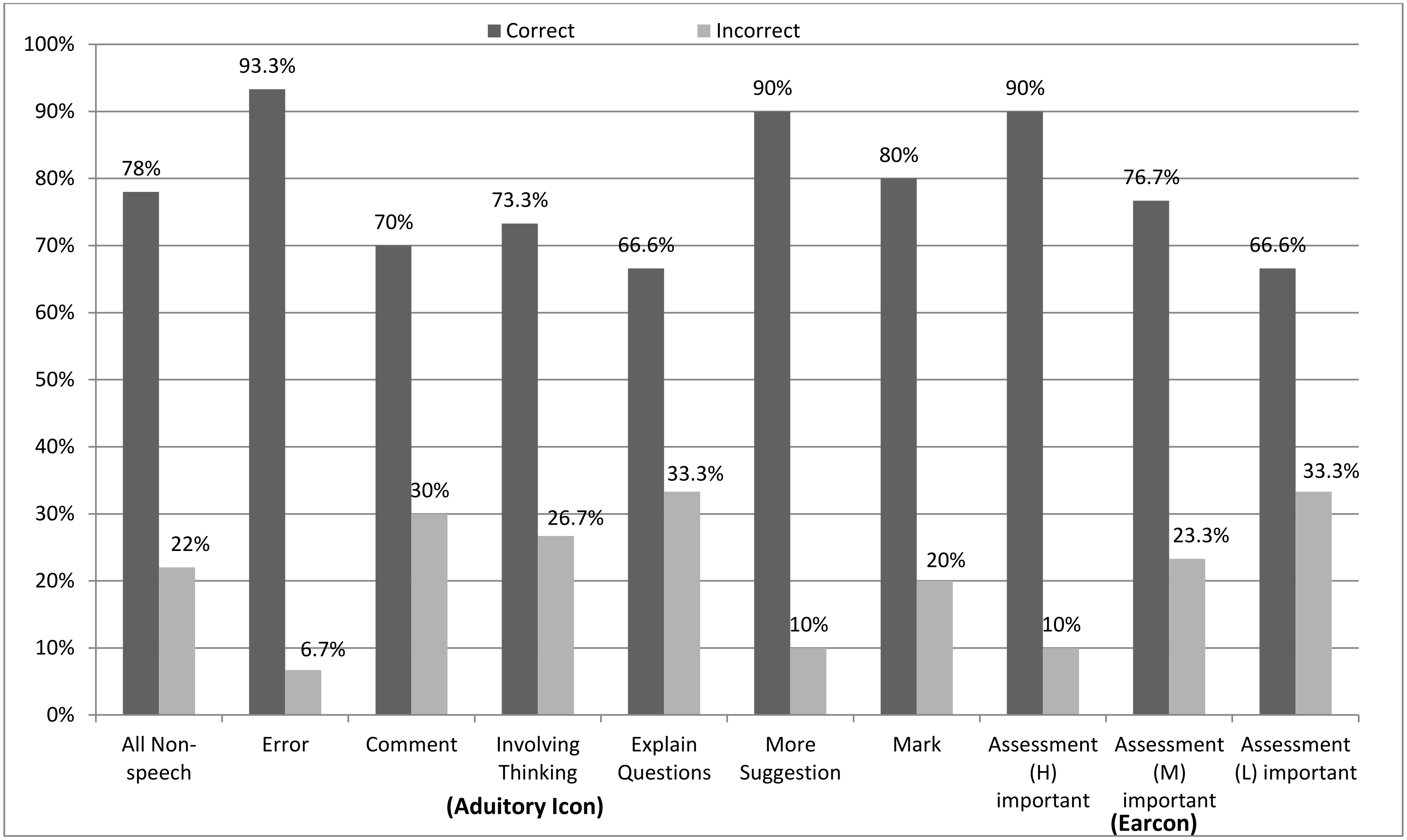



**Figure 5.4: Correct recognition of users of the assessment types that were communicated by non-speech sounds, earcons and auditory icons**

Users were requested to do three questions communicated using non-speech, metaphors to determine the high, medium or low assistive capability of the assistant types. Figure 5.4 shows that 90% (27 users) responded correctly when using a high importance assistance type compared to 76.7% (23 users) for a medium importance assistant type and 66.6% (20 users) for the low importance assistance type. The Chi-square outcomes (Table 5.3) shows that the involvement of users was increased given the correct answers as measured in the experiment. In the second involvement test 2, three sounds were played for each of the assessment types and the important level of the assistant type was determined. Users had to distinguish the sound that linked each of the assessment types and its level of importance.

| Variable | Chi-square value | Sig. | Significance |
|---|---|---|---|
| All Non-Speech | 15.6b | .004 | Yes |
| **Auditory Icon** | | | |
| Error | 4.5a | 0.031 | Yes |
| Comment | 8.5a | .003 | Yes |
| Involving Thinking | 5.8a | .003 | Yes |
| Explain Questions | 4.8a | .028 | Yes |
| More Suggestion | 22.5a | .000 | Yes |
| Mark | 6.5 a | .000 | Yes |
| **Eercons for Assessment (Ability )** | | | |
| High | $22.5^a$ | .000 | Yes |
| Medium | $19.2^a$ | .003 | Yes |
| Low | 13.3a | 068 | Yes |

**Table 5.5: Chi-square and significance calculations relating to involving test 2**

The obtained results are shown in Figure 5.5 for all non-speech sounds, earcons and auditory icons. In total, 84% of the tested sounds were correctly recognised by users. This outcome was highly significant (1) = 15.6, cv = 3.84, p < .05, see also Table 5.5) as can be seen in Figure 5.5 that 100% (30 users) correctly recognised the auditory icons broken glass for error and 93.3% opening bottle for more suggestion.

However, this percentage for the other sounds was only 76.6 % and 70% respectively. For earcons, Figure 5.5 illustrates that the sounds used for high, medium and low assistive importance were correctly recognised by all users.

The Chi-square results can be seen in in Table 5.3 and proved a significant difference among correct and incorrect recognition for all the tested sounds. In short, the obtained outcomes suggest that the tested auditory icons and earcons might be successfully interpreted and more easily remembered by users when utilised in e-assessment condition to signal the importance of particular content delivered by a body gesture.

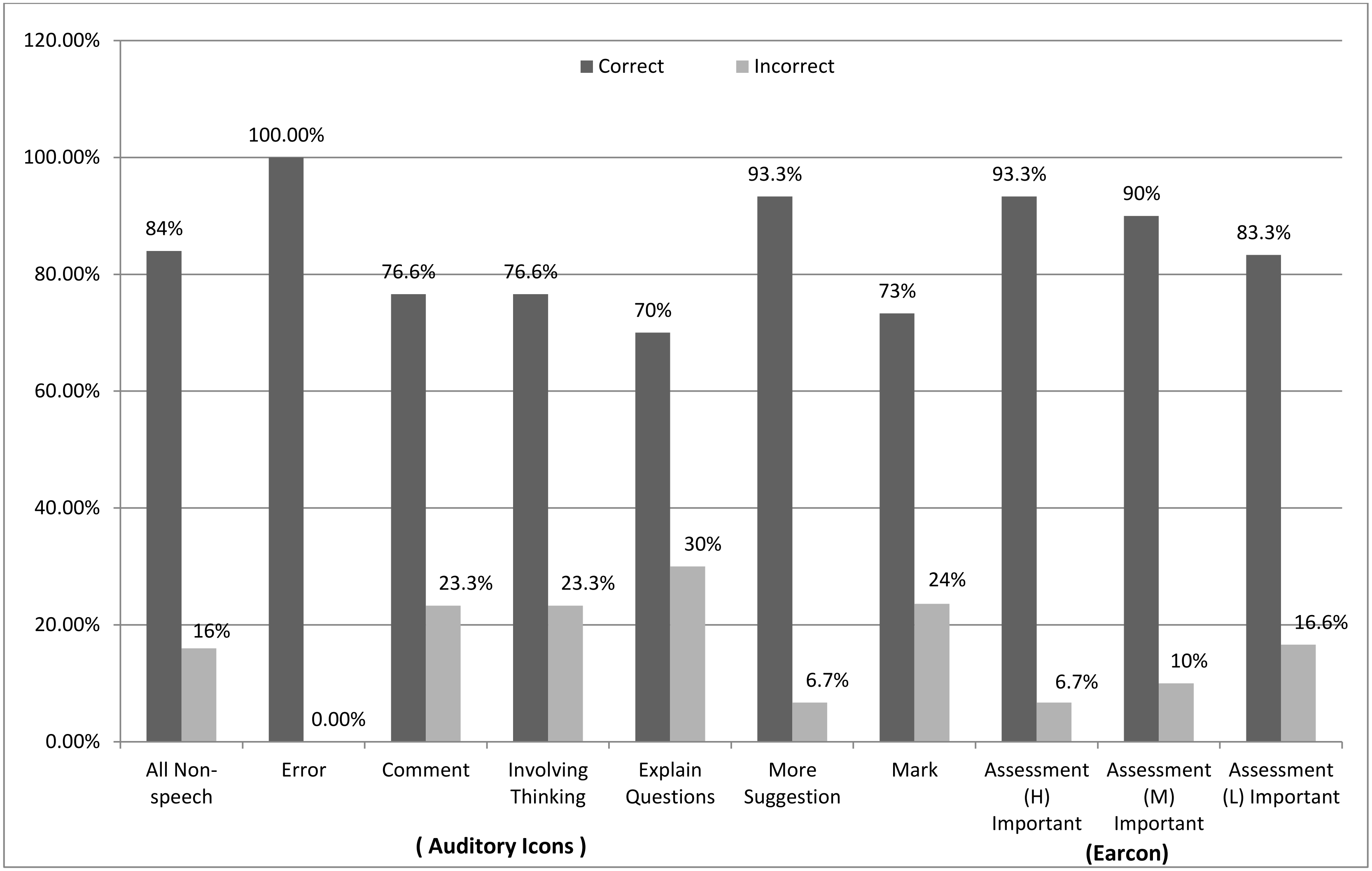



**Figure 5.5: Recognition of users for all the tested non-speech sound**

### 5.8.3 User Evaluation

Prior to the experiment, users were requested to express their views (agreement or disagreement), in terms of frustration, usefulness and concentration, based on the use of earcons and auditory icons that complemented the voice of body gesture in an e-assessment interface.

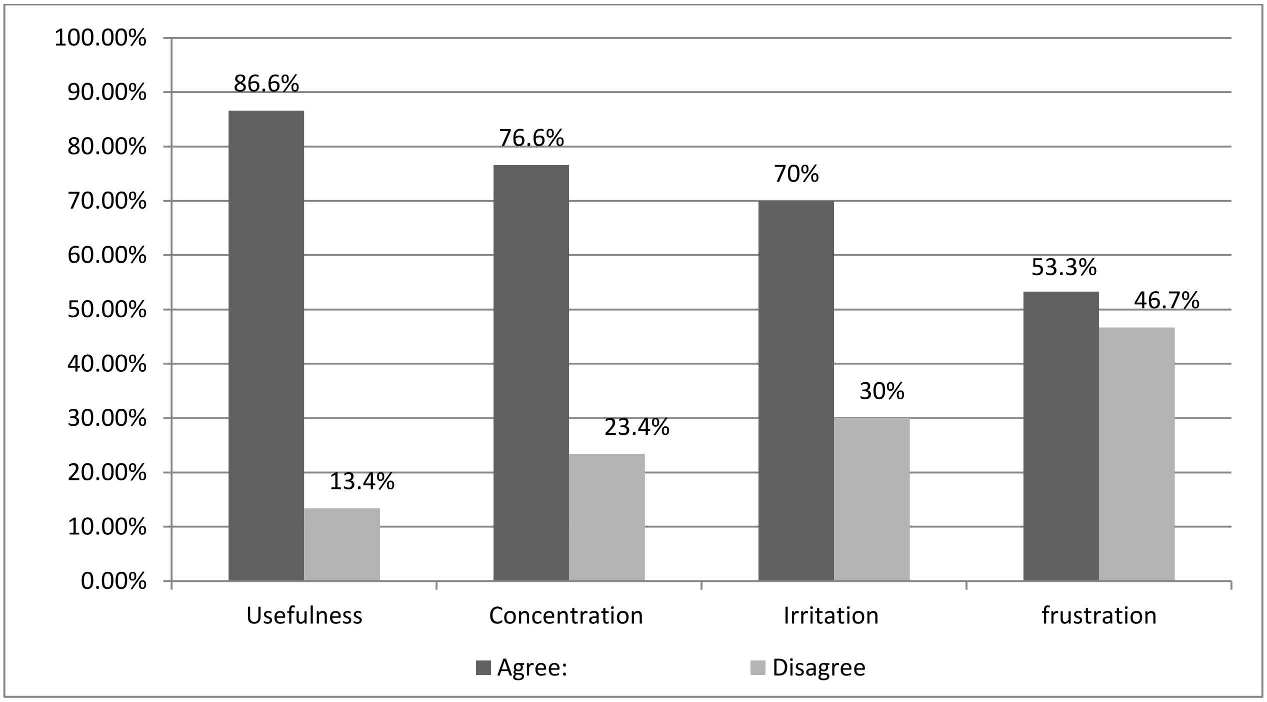



**Figure 5.6: Results of the user evaluation toward the non-speech sound**

The responses of users were positive, as shown in Figure 5.6, with respect to their views and feelings about earcons and auditory icons used interactively in the Auditory Avatar Body Gestures interface. However, 70% of the users felt irritated when they heard the sounds through the experiment. It is noteworthy that there was no large difference in frustration. In terms of student participants who felt frustration when earcons and auditory icon were offered in the interface, there was a small difference between agreement and disagreement of 53.3% and 46.7% respectively. In reference to usefulness, 86.6% of users found these sounds to be helpful. In addition, 76.6% of users felt that the presentation of sound assisted them to concentrate due to the interaction with the Auditory Avatar Body Gestures interface.

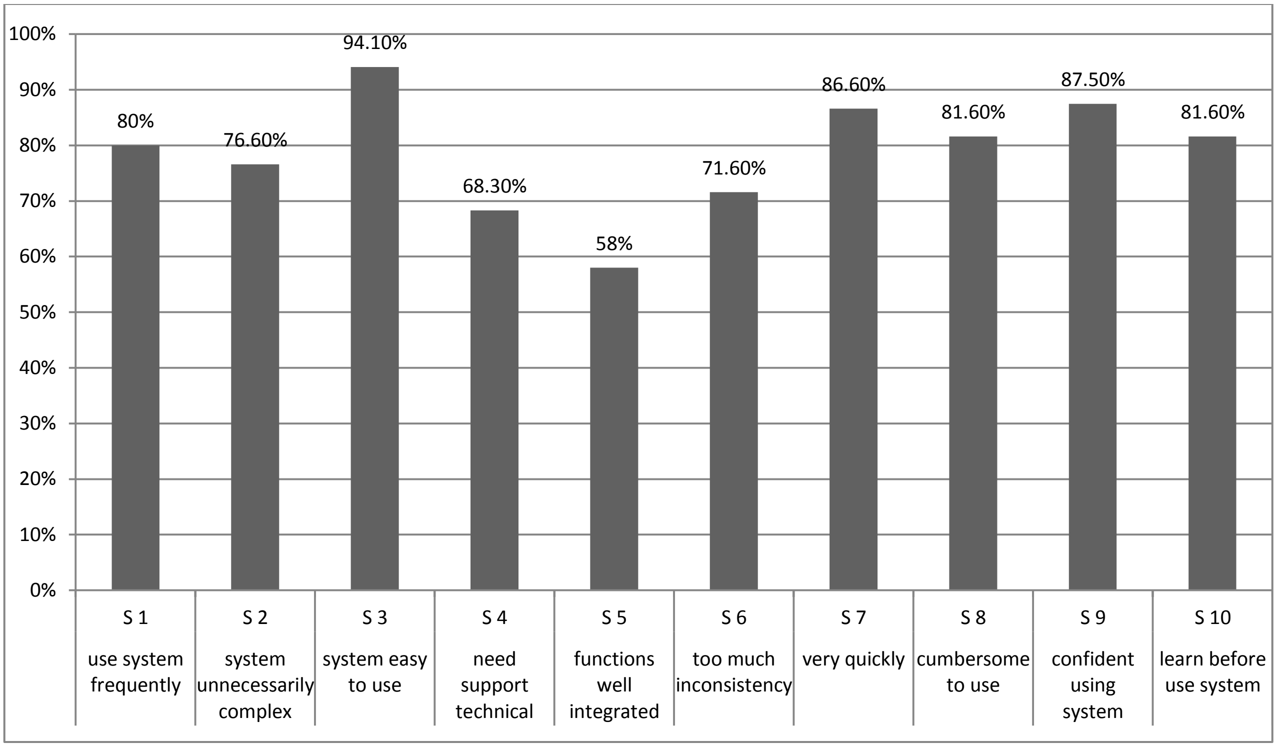



**Figure 5.7: The percentage of user responses to each statement in the satisfaction questionnaire**

### 5.8.4 Satisfaction

User satisfaction was measured using a questionnaire composed of 10 statements each of which had a 5-point Likert scale [134, 135]. Ratings for this scale were 1 to 2 were for 'Disagree' and 'Very Disagree', 3 for 'Undecided' and 4 to 5 for 'Agree and 'Very Agree' to obtain users' attitude towards the different aspects of the Auditory Avatar Body Gestures Condition.

The mean score for user satisfaction calculated using the SUS approach was 81% indicating a high level of overall satisfaction. The percentage of users' responses to each statement in the satisfaction questionnaire is shown in Figure 5.7. Most of the advantages statements (S1, S3, S5, S7, and S9) in the SUS questionnaire attained high levels of users' agreement. More topically, the condition and learning experience was satisfactory for users, 85% of the users agreed that the system functions were well integrated (S5) and that most people well learn how to use it very quickly (S7). The percentage of users who felt confident about (S9) through the interaction with the Auditory Avatar Body Gestures interface was 78.5%. The number of users who felt that they would use the system frequently was 80% and 94.1% felt that the system was simple to use. Whereas, users' disagreement regarding the disadvantages statements (S2, S4, S6, S8, and S10) was observed to be high and fluctuated between 68.2% and 81.5%. It can be noticed from these results that the majority of the users, 81.5%, need to learn before using the system (S10). On the other hand, a slightly lower percentage, 68.2%, did not agree that using the tested condition requires the need for technical support (S4).

Overall, users were excited and interested about the Auditory Avatar Body Gestures condition attributes as well as the learning materials and it was easy for the users to use. A more significant outcome was noticed in terms of a higher level of users' agreement as opposed to disagreement for the questions; this was because of the amusing earcons and auditory icons in addition to the other amusing body gestures.

## 5.9 Discussion

This experiment reveals that the users had an increased level of concentration on the delivered assessment content. This increased concentration was due to the inclusion of interaction metaphors of diverse modalities in the tested condition. The textual metaphors combined in the condition with body gestures of the assistant avatar contributed to capturing the user's visual attention towards the provided information. At the same time, additional auditory explanations about this information were presented by the voice of the full body gesture avatar.

The experiment also revealed that using non-speech sounds provided users with a mechanism to know important sign posts, for example, the ability assessment level.  Non-speech sounds did not frustrate concentration and the users found it useful (see Figure 5.6). Furthermore, users were able to be fully involved with the assessment content which was communicated by these sounds (see Figure 5.4 and 5.5). Consequently, users were able to present the correct answer as shown in Figure 5.2.

The first experimental hypothesis (H1) was concerned with the impact on achievement of the Auditory Avatar Body Gestures interface and non-speech sounds. The results of this experiment show that users' achievement levels were significantly assisted by the addition of earcons and auditory icons and aided in extending the contribution of the body animated virtual instructor to achieve both types of the required evaluation activities, namely, assistance types and recall and recognition. Auditory icons significantly assisted users' achievement level in assessment recall and recognition questions where the desire to answer these questions is attached to well recognise everyday sounds. However, earcons were more effective in recall questions than in recognition questions. The earcons used in this experiment were less helpful compared to auditory icons. The results related to users' achievement levels were significant in recall and recognition performance supporting what has been hypothesised in H1. The outcomes of the experiment indicated that the students were satisfied, significantly with the inclusion of auditory icons and earcons in the evaluated e-assessment interface (see Figure 5.6). Most users' indicated that these sounds were neither irritating nor frustrating, assisted their involvement and did not divert their concentration. Moreover, the auditory icons were chosen to assist in making a natural mapping among the assistance content types and used sounds from the immediate environment. Additionally, each of these sounds indicated one meaning at a time and was used consistently throughout the Auditory Avatar Body Gestures interface. These aspects were important particularly when they are incorporated with other auditory and visual metaphors (see 5.2, 5.3 and 5.4). This multimodality generated positive feelings with respect to the e-assessment interface. These outcomes supported all the assumptions made by the experimental hypotheses H2, H3 and H4.

Finally, the obtained results suggest that utilising non-speech sound with body gestures in the form of avatars enhances, to a large extent, the usability and user involvement within the delivery of e-assessment.

## 5.10 Summary

The experiment presented in this chapter investigated the achievement level and user involvement with the use of earcons and auditory icons used as complementary auditory signals to indicate the dissimilar assessment types as presented by a virtual instructor. The experiment also investigated users' satisfaction.

A total of 30 students took part in the experiment to assess the e-assessment interface as an extension to the interface tested in the previous experiment by adding of Non-Speech sounds. The results show that these sounds were effective in directing the users' attention to important parts of the assessment, and contributed positively to enhance user achievement levels in different learning activities. Furthermore, these sounds were memorable, understood, and increased users' satisfaction and enjoyment. Consequently, the use of these metaphors was discovered to be significantly useful to enhance the usability of an e-assessment interface. Ultimately, this chapter showed the addition of auditory non-speech metaphors to an Avatar Body Gestures condition to allow the user to engage with diverse types of Assessment and questions. Three types of multimodal metaphors were presented which were included in the interface: visual-only metaphors (text which is Assessment type content), audio-visual metaphors (speaking avatar with body gestures) and auditory metaphors (earcons and auditory icons). The collection of experimental data was mostly focussed on observations and questionnaires and contributed to the valuation of user's involvement and enhanced user performance ability, such as achievement level and user satisfaction. The results indicated that the users were satisfied, significantly with the inclusion of auditory icons and earcons. Mostly of students stated that these sounds were neither irritating nor frustrating, helped their involvement and did not divert their concentration.

The results of this study highlight the significance of the multimodal metaphors in enhancing learnability performance, as well as the usability of e-assessment interfaces, in terms of achievement level and user satisfaction.

# Chapter 6

# Conclusions and Empirically Derived Guidelines for Assessment

## 6.1 Introduction

This Chapter discusses the conclusions and empirically derived guidelines of the thesis for the incorporation of multimodal metaphors within interface structures for e-assessment applications in order to determine the usability. A brief summary of the main conclusions and empirically derived guidelines and recommendations.

## 6.2 Review of the Experimental Work

Here, a critical review of the experiments is presented. Specifically, there is a review of how the experiments were successful in achieving their aims. Overall, the aim of the experiments was to investigate the effect on usability of different types and combinations of multimodal metaphors in e-assessment. The experiment tested the use of multimodal metaphors against usability which included time taken to complete assessment successfully, overall score and satisfaction in terms of enjoyment, confidence, ease of use and independence. In order to achieve this aim, three experiments were carried out and here a review of these experiments is presented.

The aim of the first experiment was to assess the usability of different e-assessment designs that incorporated desecration text, images, and avatars. Two experimental conditions were designed for the experiment. The first condition (VOAP) was the control for the experiment and assessed usability using only text. The control was required as a measure for which the results of the experiments including multimodality were measured against. The use of the control was successful because it demonstrated usability without any multimodality; furthermore, it demonstrated a clear improvement between the control using only text and the main part of experiment one, which was the second condition (VMAP), using a multimodality including desecration text, images, and avatars.

The second condition of experiment one successfully measured efficiency using the time spent by each user to answer questions and complete tasks. Effectiveness was measured successfully using the correct answers and user satisfaction was measured using a questionnaire (chapter 3) that covered all of the predetermined areas considered to be within the scope of satisfaction. Furthermore, experiment one included a variation in terms of the difficulty of the questions and the results showed that there was a correlation between question level of difficulty and time taken and overall score.

It was expected in experiment one that users of the VMAP condition would be more satisfied than the users of the VOAP. Related to this assumption, the experiments were instrumental in determining that multimodal presentation of the learning material in the VMAP offers significantly greater satisfaction than the text only in the VOAP.

The focus of experiment two was primarily to investigate the role of images with other multimodal metaphors in improving usability. The experiment was successful because it revealed the effects on the users of different combinations of multimodal metaphors, additionally, it revealed the most effective combination in terms of usability and the opinion that users had about the e-assessment exercise. Specifically, the experiment revealed that the best combination was images with avatars and it is important to note that this result was important in informing the next experiment, experiment three, where there was a more in-depth assessment of the use of avatars with multimodal metaphors.

Both experiments one and two measured the effect of multimodal metaphors against varying difficulties in the questions, namely, easy, moderate and difficult. This was important because the study aimed to understand the effect of multimodal metaphors in assessment where varying degrees of difficulty are found. Specifically, this was important because the researcher wanted to investigate how multimodal metaphors had an effect on the user when they were answering difficult questions; this is not in terms of providing clues, but rather to provide more clarity about the question (see Chapter 5).

The third experiment was concerned with the effect on usability of using an avatar, specifically; here the avatar was a full body gesture avatar, together with earcons and auditory icons. The aim of the experiment was to determine usability as well as the opinions of the users when engaged in an interactive e-assessment context using more recently developed and sophisticated multimodal metaphors. The experiment included 30 users and the results derived from this experiment showed that the tested multimodal metaphors significantly contributed to enhancing user learning performance and the usability of e-assessment interfaces, in terms of efficiency, effectiveness and user satisfaction (see Chapter 5).

## 6.3 Limitations

The main limitation of the experiments is that they could have included more combinations of multimodal metaphors. The reason this is important is because it will extend the experimentation to more combination possibilities, thus increasing the validity of the proposed derived guidelines. Although the experiments do determine the best types of multimodal metaphors, both singular, such as images in experiment one, and in combination, such as images and avatars in experiment two and avatars (body gesture) with earcons and auditory icons in experiment three, there are obvious further combinations that be assessed.

Experiments one and two included a variable which measured usability against varying levels of difficulty in the questions, namely, easy, moderate and difficult, this was not included in experiment three and although it was an important focus of experiment three to include further variables, there was still scope to include the various levels of difficulty in the questions.

Another limitation of the study was that the participants were often in a rush to complete the e-assessment experiments, a possible solution to this limitation is to establish proper assessment conditions and consider the environmental conditions as well to replicate a real life e-assessment situation.

Finally, the use of sounds, specifically, earcons and auditory icons, were not investigated as multimodalities that could be used to help those who are visually impaired. The same was true for full body gesture avatars, where the benefits for those who are hearing impaired were not considered.

## 6.4 Derived Guidelines for the Use of Multimodality

One of the two main aims of this study was to present guidelines based on the empirical results. Specifically, the study aimed to find out the most appropriate approach to the utilisation of multimodal metaphors in e-assessment interfaces to improve usability and based on these finding a set of guidelines are presented here. These guiding principles based on the findings of this study can be used with those derived from previous studies that measured usability in terms of efficiency, effectiveness and user satisfaction.

### 6.4.1 Description Text

The findings showed that clear description text with minimal information was more effective in terms of usability. From the results the recommendation is made that description text, as a multimodal metaphor, is more effective when used alone and not with another metaphor. This recommendation is based on the fact that description text alone is more effective in terms of the user providing the correct answer (see Figure 3.5) than in combination with images (see Figure 4.5). This recommendation is also supported by the fact that description text used alone is more effective (see Figure 3.4) than used in conjunction with images in reference to the time taken to complete tasks (see Figure 4.3), in fact for the latter it was found that this combination was the least effective combination in terms of completion time.

### 6.4.2 Images

Based on the results of experiment one, it is highly recommended that designers of this type of interface should consider the use of images. The results showed that images were the most effective in terms of assessed usability, namely, effectiveness in terms of score and efficiency in terms of time taken. Because later experimentation showed that images worked well with other metaphors, it is recommended that when using a combination of metaphors that images are considered.

### 6.4.3 Avatar

Any use of avatars should be considered in combination with other metaphors. The reason for this is that avatars used alone were found to be less effective than in combination with, for example, images. More specifically, the maximization of the avatar in the e-assessment interfaces is achieved when other metaphors (textual, images or both) are placed beside the avatar in the same scene. This strategy is more useful when it is desired to bring the attention of the users to the offered information.

Consequently, the advice to designers is to create a design that causes users to base their visual and auditory attention in one area within the interface and minimise the probability of users switching their attention from one part of the interface to another. The results showed that users will be more involved when their attention is focussed in one location and efficiency and effectiveness of users decreases when their visual attention is diverted. This recommendation is also based upon other experiments that highlighted the scope of the fusing of a diverse set of information components into a single place in the interface [137].

Designers should aware of the fact that the speech given by avatars may be too time consuming which has a detrimental impact both in terms of the assessment taking too long and the user receiving too much information. In fact, in relation to the latter it was found that shorter speech by the avatar had a more beneficial impact on the user in terms of usability.

### 6.4.4 Recorded Speech with Images

Developers of e-assessment interfaces should employ natural recorded speech because it was found to increase the user's attention, however, it should be noted that this should be done in combination with images. This combination allows users to be more involved with the material. However, if the designers of interfaces are considering the use of recorded speech in combination with images, then is would be better for them to consider avatars with images instead because this combination was more effective in increasing usability. As with the avatar, it is recommended here that natural recorded speech should be used for shorter messages for communication in e-assessment, for example, to give instructions to users.

### 6.4.5 Images with Description Text

If designers wish to incorporate the use of description text in the e-assessment interface, then it is recommended that they are used alone. The reason for this is that description text alone was found to be more effective in terms of improving usability than in combination with images. Designers should note that this is the only instance where the exclusion of images is more effective.

### 6.4.6 Full Body Gestures Avatar with Earcons and Auditory Icons

It is highly recommended than designers always consider the use of an avatar in the e-assessment interface, as the results have shown this to be beneficial. However, given a choice between the avatar that is a recorded talking head with facial expressions or a full body gesture avatar, then the latter should be preferred. The reason for this was that usability was higher for the full body gesture avatar. Specifically, using full body gestures avatar performed better in terms of helping users. Furthermore, positive body gestures should be implemented in design of the full body gesture avatar in order that it may communicate more effectively areas such as error, mark and suggestion. Overall, there a positive perception from the users about the use of this type of avatar and that designers should consider the use of full body movements as it will attract the user's attention in a real face-to-face interaction. Furthermore, it is recommended that the use of a full body gesture avatar is in combination with earcons and auditory icons. The combination of speech (avatar) and non-speech (earcons and auditory icons) stimuli was found to be effective for communication in the e-assessment interface. Furthermore, this combination helps in clarifying the types and content of the assessment.

In reference to non-speech stimuli specifically, they caused users to be more engaged in the assessment and more likely to provide the correct answers. Designers should be aware that such non-specch stimuli, i.e. earcons and auditory icons, should be used in conjunction with full body gesture avatars as supporting metaphors. The results show that such a combination reduces confusion and ambiguity. Recommended examples include the sound of a bottle opening (auditory icon) to indicate that an important statement is about to commence and the sound of a window closing for the closing of a statement, and earcons can be used to annotate parts of assessment or statements. Designers should use non-speech stimuli during the pause intervals of the spoken messages by the avatar in order to avoid interference and confusion. This is important because it is more likely that the user will be able to decipher the critical parts of the assessment and have increased attention. It is important for designers to note that auditory icons were successful in assisting users in both recall and recognition questions where the required answer was communicated using auditory icons only. For earcons they were less effective in recall questions than they were in recognition questions. Furthermore, in this specific context earcons were less beneficial compared to auditory icons. In light of these findings it is recommended that designers prefer the use of auditory icons.

It is recommended that the designer uses positive expressions for example neutral, happy, and thinking because it has a significant effect on usability, and they should also note that the use of negative expressions has no impact on usability, and therefore, they should disregard expressions such as irritation and frustration.

### 6.4.7 Complexity of Assessment

Designers can be confident in using multimodal metaphors for all levels of difficulty and complexity. The result showed that the use of metaphors was more effective in terms of usability for easy, moderate and difficult questions in comparison to assessments that do not use metaphors. Moreover, the use of metaphors is highly recommended for communicating more complex ideas in the assessment and where the designers needs to improve recall and recognition methods.

## 6.5 Future Work

### 6.5.1 Smart Virtual Instruction

The avatars that were used in this study were used to communicate information related to the question, if the user needs more elaboration about the question they could ask the avatars then the avatars are capable of providing this. In relation to this idea, it is proposed that this facility is extended to include a more interactive dialogue through an avatar that has a smart facility. This could be achieved through an avatar that has a certain level if artificial intelligence which includes speech recognition technology and is connected to the content of the assessment. Example, this could include speech recognition technology. This will enable the avatar to understand the user's needs and to assist them with understanding questions accordingly.

### 6.5.2 Personalised Virtual Instructor

Although the research considered the expressions of the face and body of the avatar, it does not consider other attributes of the avatar that may also have an impact on users and usability, these could include, for example, tone and intonation of the avatar's speech or gender, culture and age of the avatar. Therefore, future work could examine these and more attributes of an avatar and how this effects the user in order to create more tailored or personalised avatars suitable for different types of user. Moreover, the users who participated in the present study were a homogenous group, i.e. students, thus a future investigation could examine the different attributes of different groups and how this would impact on the design of e-assessment interfaces.

### 6.5.3 Virtual Instructor for Less-able Users

Another area that deserves more investigation is the design of avatars and other multimodal metaphors for users with disabilities. Suggestions for this include avatars that use sign language for the hearing impaired, or the full body gestures could be used more extensively to give information, assistance or instruction. Another idea is to extend the use of earcons and auditory icons to consider those who are sight impaired.

### 6.5.4 Other Types of Assistance

The types of assistance that were investigated were related to suggestion, comments, and error. Based on the specific needs of users any future study could investigate the use of other types of assistance. This will provide a more specialised approach to communicating ideas through the e-assessment interface.

### 6.5.5 Mobile E-assessment

This research focussed on an online electronic environment. However, portable mobile computing is becoming a more popular medium and can be used for e-assessment as well. Therefore, there is a need in future studies to examine how multimodality can be exported for use in the mobile computing medium, for example, the implications of the use of multimodal metaphors on smaller screens can be investigated.

## 6.6 Conclusion

This part addresses the conclusions and limitations of the experimental outcomes in this research. The first experiment provided the fact that the usability of the experimental condition with the multimodal metaphors was better than the condition with text in offering e-assessment to users. Combining description text, avatar, and images had greater efficiency in terms of reducing the time required to answer the task and complete tasks successfully. They were helpful in enabling the student participants to respond correctly to a greater number of assessments, specifically if they were of higher complexity. Also, the percentage of the users using the multimodal e-assessments condition was better compared to the users using the text interface. This experiment nevertheless measured the total contribution of the multimodal metaphors and the combined impact in terms of usability. The second experiment with the facially expressive avatars and images showed greater efficiency, effectiveness, and user satisfaction compared to using record speech and description text with images. In addition, it was found to be empirically preferable to combine the body animated virtual tutor and the presented assessment material in one interface. In addition, the second experiment determined avatar the best in the assessed e-assessment.

The outcomes of the third experiment presented that combining earcons and auditory icons attracts user attention to the communicated assessment. The combination and synchronisation of the voice with the body gestures in the virtual tutor assisted users to involving better with the several types of activities connected to assessment. Finally, the obtained results demonstrated that users detected, comprehended, and interpreted assessment better with the aid of non-speech sounds.

Overall, when considering the results of the experiment, it can be extrapolated that it is more likely for majority of the users to enjoy (and attain satisfaction from) the multimodal assisted e-assessment experience. One of the key reasons for this is the enabling of users to complete learning tasks quicker and more accurately. Consequently, the general results found by this experimental study confirm the significance and the role of the multimodal communication metaphors, in enhancing student learning performance, as well as the usability of e-assessment interfaces, in terms of efficiency, effectiveness and user satisfaction.

The overall aim of this study was to investigate the impact on usability of multimodal metaphors. The conducted research assessed specific designs of multimodal metaphors (e.g. description text, image, avatar, full body gesture avatars, earcons and auditory icons) and their use for communication in e-assessment in interfaces, in terms of their impact on usability and the involvement of the users. Overall, the results of all the experiments showed that the use of metaphors increased usability, however, different metaphors and different combinations of metaphors had differing success in this regard. Specifically, the study showed the interplay between user's cognitive reasoning, interest, satisfaction and motivation, and multimodality in the user interface.

Currently there is shift taking place between using traditional methods for assessing learners and the introduction of e-assessment which is a relatively new approach to assessment using emerging technologies. However, there is still a need to understand the use of technology in this area, especially its appropriateness, applicability and effectiveness in the area of assessment. This study has served to address these issues for the use of multimodality in e-assessment interfaces and will be invaluable to those who are responsible for bringing together technology with assessment, namely, academics and interface designers.

These results were instrumental in the development of set guidelines for the development of e-assessment interfaces using multimodal metaphors which will be of particular use to developers and software designers. Moreover, the results contribute significantly to the research literature and offer numerous suggestions for future study to take the ideas presented here further, or to overcome present limitations.

References

[1] C. Redecker and Ø Johannessen, "Changing Assessment—Towards a New Assessment Paradigm Using ICT," *European Journal of Education,* vol. 48, pp. 79-96, 2013.

[2] J. W. Gikandi, D. Morrow and N. E. Davis, "Online formative assessment in higher education: A review of the literature," *Comput. Educ.,* vol. 57, pp. 2333-2351, 2011.

[3] M. Llamas-Nistal, M. J. Fernández-Iglesias, J. González-Tato and F. A. Mikic-Fonte, "Blended e-assessment: Migrating classical exams to the digital world," *Comput. Educ.,* vol. 62, pp. 72-87, 2013.

[4] A. Boyle and D. Hutchison, "Sophisticated tasks in e-assessment: what are they and what are their benefits?" *Assessment & Evaluation in Higher Education,* vol. 34, pp. 305-319, 2009.

[5] M. C. Johnson-Glenberg, "Embedded formative e-assessment: who benefits, who falters," *Educational Media International,* vol. 47, pp. 153-171, 2010.

[6] L. Jantschi, C. E. Stoenoiu and S. D. Bolboaca, "Linking assessment to e-learning in microbiology and toxicology for undergraduate students," in *EUROCON, 2007. the International Conference on &# 34; Computer as a Tool&# 34;* 2007, pp. 2447-2452.

[7] M. M. Lynch, *The Online Educator: A Guide to Creating the Virtual Classroom.* Routledge, 2002.

[8] H. Barbosa and F. Garcia, "Importance of online assessment in the e-learning process," in *Information Technology Based Higher Education and Training, 2005. ITHET 2005. 6th International Conference On,* 2005, pp. F3B/1-F3B/6.

[9] G. Conole and B. Warburton, "A review of computer-assisted assessment," *Research in Learning Technology,* vol. 13, 2005.

[10] J. Dermo, "e-Assessment and the student learning experience: A survey of student perceptions of e-assessment," *British Journal of Educational Technology,* vol. 40, pp. 203-214, 2009.

[11] C. N. Quinn, "Seven steps to better e-learning," *eLearn,* vol. 2006, pp. 2, 2006.

[12] C. N. Quinn, "Seven steps to better e-learning," *eLearn,* vol. 2006, pp. 2, 2006.

[13] S. Iram, D. Al-Jumeily and J. Lunn, "An integrated web-based e-assessment tool," in *Developments in E-Systems Engineering (DeSE), 2011,* 2011, pp. 271-275.

[14] G. Busuttil-Reynaud and J. Winkley, "JISC e-assessment Glossary," 2006.

[15] S. Jordan, "E-assessment: Past, present and future," *New Directions,* vol. 9, pp. 87-106, 2013.

[16] R. Stowell and R. Lamshed, "E-assessment guidelines and case studies," *Interim Report, Australian Flexible Learning Framework and National Quality Council,* 2011.

[17] N. Buzzetto-More and A. Alade, "Best practices in e-assessment," *Journal of Information Technology Education: Research,* vol. 5, pp. 251-269, 2006.

[18] H. Barbosa and F. Garcia, "Importance of online assessment in the e-learning process," in *Information Technology Based Higher Education and Training, 2005. ITHET 2005. 6th International Conference On,* 2005, pp. F3B/1-F3B/6.

[19] M. Sainsbury and T. Benton, "Designing a formative e-assessment: Latent class analysis of early reading skills," *British Journal of Educational Technology,* vol. 42, pp. 500-514, 2011.

[20] S. Ball, "Accessibility in e-assessment," *Assessment & Evaluation in Higher Education,* vol. 34, pp. 293-303, 2009.

[21] J. Dermo, "e-Assessment and the student learning experience: A survey of student perceptions of e-assessment," *British Journal of Educational Technology,* vol. 40, pp. 203-214, 2009.

[22] P. Hodgson and M. Y. Pang, "Effective formative e-assessment of student learning: a study on a statistics course," *Assessment & Evaluation in Higher Education,* vol. 37, pp. 215-225, 2012.

[23] D. Kumar Saini and L. Sunil Prakash, "Plagiarism Detection in Web based Learning Management Systems and Intellectual Property Rights in the Academic Environment." *International Journal of Computer Applications,* vol. 57, 2012.

[24] T. Wielicki, "Integrity of online testing in e-learning: Empirical study," in *Pervasive Computing and Communications Workshops, 2006. PerCom Workshops 2006. Fourth Annual IEEE International Conference On,* 2006, pp. 5 pp.-210.

[25] N. Jyothi, K. Bhan, U. Mothukuri, S. Jain and D. Jain, "A recommender system assisting instructor in building learning path for personalized learning system," in *Technology for Education (T4E), 2012 IEEE Fourth International Conference On,* 2012, pp. 228-230.

[26] Y. Caniou and T. Excoffier, "E-learning with quenlig: An online questionnaire," in *Global Engineering Education Conference (EDUCON), 2012 IEEE,* 2012, pp. 1-9.

[27] A. Dika, K. Bilali and F. Dalipi, "Implementing and using new e-testing system to increase the effectiveness of learning in primary schools of macedonia," in *Education and E-Learning Innovations (ICEELI), 2012 International Conference On,* 2012, pp. 1-4.

[28] A. L. McCann, "Factors affecting the adoption of an e-assessment system," *Assessment & Evaluation in Higher Education,* vol. 35, pp. 799-818, 2010.

[29] U. Stödberg, "A research review of e-assessment," *Assessment & Evaluation in Higher Education,* vol. 37, pp. 591-604, 2012.

[30] M. AL-Smadi, C. Gütl and V. Chang, "Addressing e-assessment practices in e-learning activities: A review," in *Global Learn,* 2011, pp. 448-459.

[31] M. Uday Kumar, J. Mamatha, S. Jain and D. Jain, "Intelligent online assessment methodology," in *Next Generation Web Services Practices (NWeSP), 2011 7th International Conference On,* 2011, pp. 215-220.

[32] N. Buzzetto-More and A. Alade, "Best practices in e-assessment," *Journal of Information Technology Education: Research,* vol. 5, pp. 251-269, 2006.

[33] N. Iahad, G. A. Dafoulas, E. Kalaitzakis and L. A. Macaulay, "Evaluation of online assessment: The role of feedback in learner-centered e-learning," in *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference On,* 2004, pp. 10 pp.

[34] M. AL-Smadi, M. Hofler and C. Guetl, "Integrated and enhanced e-assessment forms for learning: Scenarios from alice project," in *Interactive Collaborative Learning (ICL), 2011 14th International Conference On,* 2011, pp. 626-631.

[35] D. Tian, "A progress-based online assessment system for first-year networking classes," in *Frontiers in Education, 2005. FIE'05. Proceedings 35th Annual Conference,* 2005, pp. S2H-31.

[36] M. S. Crowther, C. C. Keller and G. L. Waddoups, "Improving the quality and effectiveness of computer-mediated instruction through usability evaluations," *British Journal of Educational Technology*, vol. 35, pp. 289-303, 2004.

[37] A. Koohang and J. Du Plessis, "Architecting usability properties in the e-learning instructional design process," *International Journal on E-Learning,* vol. 3, pp. 38-44, 2004.

[38] Y. Wang, "Assessment of learner satisfaction with asynchronous electronic learning systems," *Information & Management,* vol. 41, pp. 75-86, 2003.

[39] R. D. Johnson, S. Hornik and E. Salas, "An empirical examination of factors contributing to the creation of successful e-learning environments," *International Journal of Human-Computer Studies,* vol. 66, pp. 356-369, 2008.

[40] R. P. Ferretti and C. M. Okolo, "Authenticity in learning: multimedia design projects in the social studies for students with disabilities," *J. Learn. Disabil.*, vol. 29, pp. 450-460, Sep, 1996.

[41] B. Hansen, "The Dictionary ofMultimedia, Terms and Acronyms," 1997.

[42] C. Miller, "Aesthetics and e-assessment: the interplay of emotional design and learner performance," *Distance Education,* vol. 32, pp. 307-337, 2011.

[43] B. Mason and R. Bruning, *Providing Feedback in Computer-Based Instruction: What the Research Tells Us, 2001,* 2010.

[44] T. Miller, "Formative computer-based assessment in higher education: the effectiveness of feedback in supporting student learning," *Assessment & Evaluation in Higher Education,* vol. 34, pp. 181-192, 2009.

[45] H. Wang, M. Chignell and M. Ishizuka, "Improving the usability and effectiveness of online learning: How can avatars help?" in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* 2005, pp. 769-773.

[46] J. Rubin and D. Chisnell, *Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests.* John Wiley & Sons, 2008.

[47] A. Gibbons, *Multimodality, Cognition, and Experimental Literature.* Routledge, 2012.

[48] C. Raymaekers, "Special issue on enactive interfaces," *Interact Comput,* vol. 21, pp. 1-2, 2009.

[49] S. Krug, "Don’t make me think: A common sense approach to Web usability: How we really use the Web," *Retrieved April,* 2009.

[50] A. Dix, J. Finlay and G. Abowd, "D & BEALE, R.(2004)," *Hum. -Comput. Interact.,* .

[51] Z. Obrenovic and D. Starcevic, "Modeling multimodal human-computer interaction," *Computer,* vol. 37, pp. 65-72, 2004.

[52] A. Jaimes and N. Sebe, "Multimodal human–computer interaction: A survey," *Comput. Vision Image Understanding,* vol. 108, pp. 116-134, 2007.

[53] L. Barfield, *Design for New Media: Interaction Design for Multimedia and the Web.* Pearson Education, 2004.

[54] L. Brown, S. Brewster, R. Ramloll, W. Yu and B. Riedel, "Browsing modes for exploring sonified line graphs," in *IN PROCEEDINGS OF BCS-HCI 2002,* 2002, .

[55] N. B. Sarter, "Multimodal information presentation: Design guidance and research challenges," *Int. J. Ind. Ergonomics,* vol. 36, pp. 439-445, 2006.

[56] S. A. Brewster, "Using non-speech sound to overcome information overload," *Displays,* vol. 17, pp. 179-189, 1997.

[57] M. M. Blattner, D. A. Sumikawa and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human–Computer Interaction,* vol. 4, pp. 11-44, 1989.

[58] A. Jaimes and N. Sebe, "Multimodal human–computer interaction: A survey," *Comput. Vision Image Understanding,* vol. 108, pp. 116-134, 2007.

[59] L. Barfield, *Design for New Media: Interaction Design for Multimedia and the Web.* Pearson Education, 2004.

[60] K. Lyon and P. J. Nürnberg, "Multimodal modelling languages to reduce visual overload in UML diagrams," in *Program Visualization Workshop,* 2004, pp. 41.

[61] (). *RefGrab-It Install Page*.

[62] S. Brewster and L. M. Brown, "Tactons: Structured tactile messages for non-visual information display," in *Proceedings of the Fifth Conference on Australasian User Interface-Volume 28,* 2004, pp. 15-23.

[63] S. Brewster, *Providing a Model for the use of Sound in User Interfaces.* University of York, Department of Computer Science, 1992.

[64] B. H. Deatherage, "Auditory and other sensory forms of information presentation," *Human Engineering Guide to Equipment Design,* pp. 123-160, 1972.

[65] M. S. Sanders and E. J. McCormick, *Human Factors in Engineering and Design* . McGRAW-HILL book company, 1987.

[66] S. J. Mountford and W. Gaver, "Talking and listening to computers," *The Art of Human-Computer Interface Design,* pp. 319-334, 1990.

[67] W. W. Gaver, "The SonicFinder: An interface that uses auditory icons," *Hum. -Comput. Interact.,* vol. 4, pp. 67-94, 1989.

[68] S. Lohr and J. Markoff, "Computers learn to listen, and some talk back," *New York Times,* 2010.

[69] D. O. Johnson and A. Agah, "Human robot interaction through semantic integration of multiple modalities, dialog management, and contexts," *International Journal of Social Robotics,* vol. 1, pp. 283-305, 2009.

[70] W. W. Gaver, "Auditory icons: Using sound in computer interfaces," *Hum. -Comput. Interact.,* vol. 2, pp. 167-177, 1986.

[71] M. M. Blattner, D. A. Sumikawa and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human–Computer Interaction,* vol. 4, pp. 11-44, 1989.

[72] G. Kramer, "An introduction to auditory display," in *Auditory Display: Sonification, Audification and Auditory Interfaces, SFI Studies in the Sciences of Complexity, Proceedings,* 1994, pp. 1-77.

[73] W. Buxton, "Introduction to this special issue on nonspeech audio," *Human Computer Interaction,* vol. 4, pp. 1-9, 1989.

[74] M. B. Alotaibi and D. I. Rigas, "A usability evaluation of multimodal metaphors for customer knowledge management," *International Journal of Computers and Communications,* vol. 2, 2008.

[75] M. B. Alotaibi and D. I. Rigas, "An empirical approach to multimodal customer knowledge management," *Intelligent Decision Technologies,* vol. 4, pp. 181-195, 2010.

[76] M. Alotaibi and D. Rigas, "The role of avatars with facial expressions to communicate customer knowledge," *International Journal of Computers, NAUN,* vol. 3, pp. 1-10, 2009.

[77] B. N. Walker, A. Nance and J. Lindsay, "Spearcons: Speech-based earcons improve navigation performance in auditory menus," in *Proceedings of the International Conference on Auditory Display, London, UK,* 2006, pp. 63-68.

[78] W. W. Gaver, "Auditory interfaces," *Handbook of Human-Computer Interaction,* vol. 1, pp. 1003-1041, 1997.

[79] E. Brazil, M. Fernström and J. Bowers, "Exploring concurrent auditory icon recognition," in *Proc. of ICAD,* 2009, .

[80] P. Polotti and D. Rocchesso, *Sound to Sense, Sense to Sound: A State of the Art in Sound and Music Computing.* Logos Verlag Berlin GmbH, 2008.

[81] W. W. Gaver, R. B. Smith and T. O'Shea, "Effective sounds in complex systems: The ARKola simulation," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 1991, pp. 85-90.

[82] M. G. Helander, T. K. Landauer and P. V. Prabhu, *Handbook of Human-Computer Interaction.* Elsevier, 1997.

[83] S. Conversy, "Ad-hoc synthesis of auditory icons," *Proceedings of of ICAD,* 1998.

[84] K. Tuuri, M. Mustonen and A. Pirhonen, "Same sound–different meanings: A novel scheme for modes of listening," *Proceedings of Audio Mostly,* pp. 13-18, 2007.

[85] A. Darvishi, E. Munteanu, V. Guggiana, H. Schauer, M. Motavalli and M. Rauterberg, "Designing environmental sounds based on the results of interaction between objects in the real world," in *Human-Computer Interaction–INTERACT,* 1995, pp. 38-42.

[86] M. A. Garcia-Ruiz and J. R. Gutierrez-Pulido, "An overview of auditory display to assist comprehension of molecular information," *Interact Comput,* vol. 18, pp. 853-868, 2006.

[87] S. A. Brewster, "Using earcons to improve the usability of tool palettes," in *CHI 98 Cconference Summary on Human Factors in Computing Systems,* 1998, pp. 297-298.

[88] M. N. A. Mazlan and L. Burd, "Does an avatar motivate?" in *Frontiers in Education Conference (FIE), 2011,* 2011, pp. T4J-1-T4J-6.

[89] R. Sheth, "Avatar technology: giving a face to the e-learning interface," *The eLearning Developers' Journal,* pp. 1-10, 2003.

[90] H. Choi, "Social learning through the avatar in the virtual world: The effect of experience type and personality type on achievement motivation," in *Society for Information Technology & Teacher Education International Conference,* 2010, pp. 1866-1873.

[91] Y. Jang, W. Kim and S. Ryu, "An exploratory study on avatar-self similarity, mastery experience and self-efficacy in games," in *Advanced Communication Technology (ICACT), 2010 the 12th International Conference On,* 2010, pp. 1681-1684.

[92] T. McLaughlin, D. Smith and I. A. Brown, "A framework for evidence based visual style development for serious games," in *Proceedings of the Fifth International Conference on the Foundations of Digital Games,* 2010, pp. 132-138.

[93] H. Wang, M. Chignell and M. Ishizuka, "Improving the usability and effectiveness of online learning: How can avatars help?" in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* 2005, pp. 769-773.

[94] G. Castellano, G. Caridakis, A. Camurri, K. Karpouzis, G. Volpe and S. Kollias, "Body gesture and facial expression analysis for automatic affect recognition," *Blueprint for Affective Computing: A Sourcebook,* pp. 245-255, 2010.

[95] B. De Gelder, "Towards the neurobiology of emotional body language," *Nature Reviews Neuroscience,* vol. 7, pp. 242-249, 2006.

[96] N. Gazepidis and D. Rigas, "Evaluation of Facial Expressions and Body Gestures in Interactive Systems," *NAUN International Journal of Computers,* vol. 1, pp. 92-97, 2008.

[97] K. Jokinen and A. Raike, "Multimodality–technology, visions and demands for the future," in *Proceedings of the 1st Nordic Symposium on Multimodal Interfaces*, 2003, pp. 239-251.

[98] C. Baber and B. Mellor, "Using critical path analysis to model multimodal human–computer interaction," *International Journal of Human-Computer Studies,* vol. 54, pp. 613-636, 2001.

[99] N. B. Sarter, "Multimodal information presentation: Design guidance and research challenges," *Int. J. Ind. Ergonomics,* vol. 36, pp. 439-445, 2006.

[100] D. Rigas and M. Alsuraihi, "A toolkit for multimodal interface design: An empirical investigation," in *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*Anonymous Springer, 2007, pp. 196-205.

[101] D. Rigas and N. Gazepidis, "A further investigation of facial expressions and body gestures as metaphors in e-commerce," in *International Conference on Applied Informatics and Communications,* 2007, pp. 148-153.

[102] S. Buisine and J. Martin, "Design principles for cooperation between modalities in bi-directional multimodal interfaces," in *Proceedings of the CHI 2003 Workshop on Principles for Multimodal User Interface Design, Ft. Lauderdale, Florida,* 2003, .

[103] M. Alseid and D. Rigas, "Efficiency of multimodal metaphors in the presentation of learning information." in *BCS HCI (2),* 2008, pp. 107-110.

[104] A. Ciuffreda and D. Rigas, "A usability Study of multimodal interfaces for the presentation of Internet Search Results," *International Journal of Computers,* vol. 2, pp. 120-125, 2008.

[105] S. A. Brewster, P. C. Wright and A. D. Edwards, "The design and evaluation of an auditory-enhanced scrollbar," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 1994, pp. 173-179.

[106] B. Duggan and M. Deegan, "Considerations in the usage of text to speech (tts) in the creation of natural sounding voice enabled web systems," in *Proceedings of the 1st International Symposium on Information and Communication Technologies*, 2003, pp. 433-438.

[107] A. Catania, "Learning-Remembering," 1992.

[108] D. K. McGookin and S. A. Brewster, "Empirically derived guidelines for the presentation of concurrent earcons," *Proceedings of BCS-HCI 2004,* vol. 2, pp. 65-68, 2004.

[109] M. Turk, "Multimodal interaction: A review," *Pattern Recog. Lett.,* vol. 36, pp. 189-195, 2014.

[110] A. A. Alharbi, "Multimodal e-feedback: An Empirical Study," 2013.

[111] T. Stivers and J. Sidnell, "Introduction: multimodal interaction," *Semiotica,* vol. 2005, pp. 1-20, 2005.

[112] B. Dumas, D. Lalanne and S. Oviatt, "Multimodal interfaces: A survey of principles, models and frameworks," in *Human Machine Interaction*Anonymous Springer, 2009, pp. 3-26.

[113] J. Lee and C. Spence, "Assessing the benefits of multimodal feedback on dual-task performance under demanding conditions," in *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction-Volume 1,* 2008, pp. 185-192.

[114] N. B. Sarter, "Multimodal information presentation: Design guidance and research challenges," *Int. J. Ind. Ergonomics,* vol. 36, pp. 439-445, 2006.

[115] L. M. Reeves, J. Lai, J. A. Larson, S. Oviatt, T. Balaji, S. Buisine, P. Collings, P. Cohen, B. Kraal and J. Martin, "Guidelines for multimodal user interface design," *Commun ACM,* vol. 47, pp. 57-59, 2004.

[116] M. M. Blattner, D. A. Sumikawa and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human–Computer Interaction,* vol. 4, pp. 11-44, 1989.

[117] M. Fabri, D. J. Moore and D. J. Hobbs, "Face value: Towards emotionally expressive avatars," *AISB02, Imperial College, London, UK,* 2002.

[118] J. Holmes, "Designing agents to support learning by explaining," *Comput. Educ.,* vol. 48, pp. 523-547, 2007.

[119] L. A. Annetta and S. Holmes, "Creating presence and community in a synchronous virtual learning environment using avatars," *International Journal of Instructional Technology and Distance Learning,* vol. 3, pp. 27-43, 2006.

[120] C. Bauckhage, T. Kaster, M. Pfeiffer and G. Sagerer, "Content-based image retrieval by multimodal interaction," in *Industrial Electronics Society, 2003. IECON'03. the 29th Annual Conference of the IEEE,* 2003, pp. 1882-1887.

[121] A. Field, "Repeated-measures designs," *Discovering Statistics using SPSS,* pp. 427-482, 2005.

[122] J. Brooke, "SUS-A quick and dirty usability scale," *Usability Evaluation in Industry,* vol. 189, pp. 194, 1996.

[123] H. Tabbers, L. Kester, H. Hummel and R. Nadolski, "Interface design for digital courses," *Integrated E-Learning: Implications for Pedagogy, Technology and Organization,* pp. 100-111, 2004.

[124] J. Moursund, J. H. Kranzler and G. Kranzler, *Statistics for the Terrified.* Prentice Hall, 2006.

[125] J. Brooke, "SUS-A quick and dirty usability scale," *Usability Evaluation in Industry,* vol. 189, pp. 194, 1996.

[126] G. Argyrous, *Statistics for Research: With a Guide to SPSS.* Sage Publications, 2011.

[127] A. Field, *Discovering Statistics using SPSS.* Sage publications, 2009.

[128] R. Sheth, "Avatar technology: giving a face to the e-learning interface," *The eLearning Developers' Journal,* pp. 1-10, 2003.

[129] S. Alharbi and D. Rigas, "Email threads: A comparative evaluation of textual, graphical and multimodal approaches," *International Journal of Computers,* vol. 3, 2009.

[130] M. Fabri and D. Moore, "Is empathy the key? effective communication via instant messaging," in *Proceedings of the 11th EATA International Conference on Networking Entities,* 2005, .

[131] J. L. Alty and D. Rigas, "Exploring the use of structured musical stimuli to communicate simple diagrams: the role of context," *International Journal of Human-Computer Studies,* vol. 62, pp. 21-40, 2005.

[132] L. M. Brown and S. A. Brewster, "Drawing by ear: Interpreting sonified line graphs," in 2003, .

[133] D. I. Rigas, *Guidelines for Auditory Interface Design: An Empirical Investigation.,* 1996.

[134] I. Nix and A. Wyllie, "Exploring design features to enhance computer-based assessment: Learners' views on using a confidence-indicator tool and computer-based feedback," *British Journal of Educational Technology,* vol. 42, pp. 101-112, 2011.

[135] B. Price and M. Petre, "Teaching programming through paperless assignments: An empirical evaluation of instructor feedback," in *ACM SIGCSE Bulletin,* 1997, pp. 94-99.

[136] S. Lohr and J. Markoff, "Computers learn to listen, and some talk back," *New York Times,* 2010.

[137] M. S. Sanders and E. J. McCormick, *Human Factors in Engineering and Design* . McGRAW-HILL book company, 1987.

# Appendix

## Appendix A – Experimental Phase 1: The Role of Multimodal E-assessment Interfaces

### A.1.1: Experimental Group.

I am happy to introduce myself to you as one of the postgraduate research students in the School of Technology at the University of De Montfort fort. I am at present investigating the use of multimodal metaphors in e-assessment interfaces, and I would like to find your views about the use of multimodal metaphors such as: Avatar, also visual text and images. Please follow the next procedure:

 Answer the pre-experiment questions.

 Read every task carefully.

 Execute the task.

 In conclusion of the task, answer the essential link questions.

 After that completion of every task, answer the satisfaction questions.

Please answer every part of the question as honestly as possible. It would be appreciative if you could fill in the next questionnaire sincerely and give your views. Thank you and me greatly your participation.

### Part 1 Pre- experiment

1- What is your age?

 18 - 24.  25 - 34.  35 - 44.  45 - 54.  55 +.

2- What is you gender?

 Male  Female

3- What is your education stage?

□ High School □ College □ Under-graduate □ Master Degree □ Doctorate Degree □ Other Area of study: ……………………………………………………………..

4- How usually do you use the computer (rate) per week?

□ 1-5 hours □ 6-10 hours □ More than 10 hours

5- Did you perform the use of any e-learning web sites or software?

□ Yes □ No

6- Do you have knowledge about Multimodal?

□ Yes □ No

7- What are the major reasons you use the Internet?

□ Surveying □ Email □ Education

8- Do you think the exam in e-assessment is good way to enhance E-learning?

□ Yes □ No

**"Input interface"**

**Task 1:**

- Please open program www.ebookexam.com
- enter your name and number in the box
- enter login in the task are coming in the first page
- choose exam 1

- **Task 2** :
- To achieve this task, perform the following requirements and answer the questions:
- Move the mouse cursor over the "question".
- Move the mouse cursor over the True or False of - the answer.

• Move the mouse over the arrow the answer and press now on the right answer T or F please the following question; show the text it might be help you to get a right answer.

1- Print is machine passes information from the user to the computer (    )

- Move the mouse cursor over the “Next to page ".

**Task 3 :** To achieve this task, perform the following requirements and answer the questions :

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following question, but before press the answer press on the Avatar might be help you to get a right answer.

2- which one of the following is a water soluble vitamin :

- o Vitamin C
- o Vitamin D
- o Vitamin K

- Move the mouse cursor over the “Next to page ".

**Task 4**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following question, show the image might be help you to get a right answer.

3- The most abundant element in the earth's crust is :

- Silicon
- Nitrogen
- Oxygen

- Move the mouse cursor over the “Next to page ".

**Task 5**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following question, show the text it might be help you to get a right answer.

4- Multimedia is :

- text, graphics, sound, and/ video
- Virtual environment
- Both of them

- Move the mouse cursor over the “Next to page ".

**Task 6:** To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following questions, but before press the answer press on the Avatar might be help you to get a right answer.

5- Thalassemia is a hereditary disease affecting:

- Blood
- Heart
- Kidney

- Move the mouse cursor over the “Next to page ".

**Task 7 :** To achieve this task, perform the following requirements and answer the questions :

• Move the mouse cursor over the "question".

• Move the mouse cursor over the True or Foals of - the answer.

• Move the mouse over the arrow the answer and press now on the right answer T or F please the following question, show the image might be help you to get a right answer.

6- Mosaic is the decorative art of creating pictures and patterns on a surface by setting small coloured pieces of glass (     )

- Move the mouse cursor over the “Next to page ".

**Task 8**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following questions, show the text it might be help you to get a right answer.

20 % of 2 are equal to:

A. 20

B. 4

C. 0.4

D. 0.04

- Move the mouse cursor over the “Next to page ".

**Task 9**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the "question".

- Move the mouse cursor over the True or Foals of **-** the answer.

- Move the mouse over the arrow the answer and press now on the right answer T or F please the following questions, but before press the answer press on the Avatar might be help you to get a right answer.

8- Australia was discovered by James Cook ( ).

- Move the mouse cursor over the “Next to page ".

**Task 10**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of **-** the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following question, show the image might be help you to get a right answer.

9- Famous graffiti artists by:

- o A master- Antoni Gaudi
- o Jean-Michel Basquiat
- o Leonardo da Vinci

- Move the mouse cursor over the “Finish “to show your time and your score and click to logout.

## Satisfaction Questionnaire

For each statement below, please express your view by placing a circle in the appropriate column.

**5**= **(SA)** Strongly Agree. **4**= **(A)** Agree. **3**= **(N)** Natural. **2**= **(D)** Disagree. **1**= **(SD)** Strongly Disagree.

| | Statement | (SA) | (A) | (N) | (D) | (SD) |
|---|---|---|---|---|---|---|
| S1 | I think that I would like to use this system frequently | 5 | 4 | 3 | 2 | 1 |
| S2 | I found the system unnecessarily complex | 5 | 4 | 3 | 2 | 1 |
| S3 | I thought the system was easy to use | 5 | 4 | 3 | 2 | 1 |
| S4 | I think that I would need the support of a technical person to be able to use this system | 5 | 4 | 3 | 2 | 1 |
| S5 | I found the various functions in this system were well integrated | 5 | 4 | 3 | 2 | 1 |
| S6 | I thought there was too much inconsistency in this system | 5 | 4 | 3 | 2 | 1 |
| S7 | I would imagine that most people would learn to use this system very quickly | 5 | 4 | 3 | 2 | 1 |
| S8 | I found the system very cumbersome to use | 5 | 4 | 3 | 2 | 1 |
| S9 | I felt very confident using the system | 5 | 4 | 3 | 2 | 1 |
| S10 | I needed to learn a lot of things before I could get going with this system | 5 | 4 | 3 | 2 | 1 |

If you have any comments or suggestions can you write please.

.............................................................................................................................................................

.........................................................................................................................................................

Thank you for your help

### A.1.2: Control Group

I am happy to introduce myself to you as one of the postgraduate research students in the School of Technology at the University of De Montfort. I am at present investigating the use of multimodal metaphors in e-assessment interfaces, and I would like to find your views about the use of multimodal metaphors such as: Avatar, also visual text and images. Please follow the next procedure:

□ Answer the pre-experiment questions.

□ Read every task carefully.

□ Execute the task.

□ In conclusion of the task, answer the essential link questions.

□ After that completion of every task, answer the satisfaction questions.

Please answer every part of the question as honestly as possible. It would be appreciative if you could fill in the next questionnaire sincerely and give your views. Thank you and I greatly your participation.

**Part 1 Pre- experiment**

1- What is your age?

□ 18 - 24. □ 25 - 34. □ 35 - 44. □ 45 - 54. □ 55 +.

2- What is your gender?

□ Male □ Female

3- What is your education stage?

□ High School (College, undergraduate, Master Degree, Doctorate Degree, or Other Area of Study) : ……………………………………………………..

4- How usually do you use the computer (rate) per week?

□ 1-5 hours □ 6-10 hours □ More than 10 hours

5- Did you perform the use of any e-learning web sites or software?

□ Yes □ No

6- Do you have knowledge about Multimodal?

 Yes  No

7- What are the major reasons you use the Internet?

 Surveying  Email  Education

8- Do you think the exam in e-Assessment is good way to enhance E-learning?

 Yes  No

**“Input interface”**

**Task 1:**

- Please open program www.ebookexam.com
- Enter your name and number in the box
- Enter login in the task are coming in the first page
- Choose assessment 1

**Task 2**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the "question".
- Move the mouse cursor over the True or False of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer T or F please the following question.

5- Print is machine passes information from the user of the computer ()

- Move the mouse cursor over the “Next to page ".

**Task 3:** To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.

- Move the mouse over the arrow the answer and press now on the right answer please the following question.

6- Which one of the following is a water soluble vitamin:

- ○ Vitamin C
- ○ Vitamin D
- ○ Vitamin K

- Move the mouse cursor over the “Next to page ".

**Task 4**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following question.

7- The most abundant element in the earth's crust is :

- ○ Silicon
- ○ Nitrogen
- ○ Oxygen

- Move the mouse cursor over the “Next to page ".

**Task 5**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following question.

8- Multimedia is :

- ○ Text, graphics, sound, and/ video
- ○ Virtual environment
- ○ Both of them

- Move the mouse cursor over the “Next to page ".

**Task 6**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following questions.

5- Thalassaemia is a hereditary disease affecting:

- ○ Blood
- ○ Heart
- ○ Kidney

- Move the mouse cursor over the “Next to page ".

**Task 7**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the "question".
- Move the mouse cursor over the True or Foals of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer T or F please the following question.

7- Mosaic is the decorative art of creating pictures and patterns on a surface by setting small coloured pieces of glass (    )

- Move the mouse cursor over the “Next to page ".

**Task 8**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following questions.

8- 20 % of 2 is equal to:

A. 20

B. 4

C. 0.4

D. 0.04

- Move the mouse cursor over the “Next to page ".

**Task 9**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the "question".
- Move the mouse cursor over the True or Foals of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer T or F please the following questions, but before press the answer press on the Avatar might be help you to get a right answer.

9- Australia was discovered by James Cook (    ) .

- Move the mouse cursor over the " Next to page ".

**Task 10**: To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the “Next to page ".
- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following question.

10- Famous graffiti artists by:

- o A master- Antoni Gaudi
- o Jean-Michel Basquiat
- o Leonardo da Vinci

- Move the mouse cursor over the “Finish “to show your time and your score and click to logout.

## Satisfaction Questionnaire

For each statement below, please express your view by placing a circle in the appropriate column.

**5**= **(SA)** Strongly Agree. **4**= **(A)** Agree. **3**= **(U)** Undecided. **2**= **(D)** Disagree. **1**= **(SD)** Strongly Disagree.

| Statement | | (SA) | (A) | (N) | (D) | (SD) |
|---|---|---|---|---|---|---|
| S1 | I think that I would like to use this system frequently | 5 | 4 | 3 | 2 | 1 |
| S2 | I found the system unnecessarily complex | 5 | 4 | 3 | 2 | 1 |
| S3 | I thought the system was easy to use | 5 | 4 | 3 | 2 | 1 |
| S4 | I think that I would need the support of a technical person to be able to use this system | 5 | 4 | 3 | 2 | 1 |
| S5 | I found the various functions in this system were well integrated | 5 | 4 | 3 | 2 | 1 |
| S6 | I thought there was too much inconsistency in this system | 5 | 4 | 3 | 2 | 1 |
| S7 | I would imagine that most people would learn to use this system very quickly | 5 | 4 | 3 | 2 | 1 |
| S8 | I found the system very cumbersome to use | 5 | 4 | 3 | 2 | 1 |
| S9 | I felt very confident using the system | 5 | 4 | 3 | 2 | 1 |
| S10 | I needed to learn a lot of things before I could get going with this system | 5 | 4 | 3 | 2 | 1 |

If you have any comments or suggestions can you write please.

..............................................................................................................................................

Thank you for your help.

**A2: User Profile Data:**

| | | Experiment group | Control Group |
|---|---|---|---|
| Age | 18- 24 | 80% | 60% |
| | 25- 30 | 20% | 40% |
| | 31- 40 | | |
| Gender | Male | 40% | 60% |
| | Female | 60% | 40% |
| Education level | Undergraduate level | 73.3% | 67% |
| | Postgraduate level | 26.7% | 33% |
| Use of computer | 1 – 5 hours | 26.7% | 13.3% |
| | 6 – 10 hours | 26.7% | 40% |
| | More than 10 hours | 46.7% | 46.7% |
| E-learning experience | Yes | 53.3% | 66.7% |
| | No | 46.7% | 33.3% |
| Multimodal Knowledge | Yes | 66.7% | 53.3% |
| | No | 33.3% | 46.7% |
| Internet use reason | Surveying | 60% | 53.3% |
| | Email | 13.3% | 20% |
| | Education | 26.7% | 26.7% |
| E-assessment enhance E-learning | Yes | 80% | 60% |
| | No | 40% | 40% |

**A3: Raw data of time spent to answer questions in Experimental group.**

| User | Level complexity | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Easy task | | | Moderate task | | | Complexity task | | |
| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 |
| U1 | 10 | 20 | 8 | 27 | 30 | 28 | 40 | 45 | 38 |
| U2 | 12 | 23 | 10 | 28 | 35 | 17 | 35 | 45 | 30 |
| U3 | 11 | 27 | 10 | 22 | 25 | 20 | 33 | 46 | 31 |
| U4 | 11 | 22 | 10 | 15 | 23 | 12 | 22 | 30 | 24 |
| U5 | 9 | 25 | 8 | 20 | 27 | 15 | 17 | 34 | 20 |
| U6 | 12 | 15 | 9 | 18 | 22 | 19 | 23 | 39 | 22 |
| U7 | 13 | 20 | 12 | 17 | 30 | 14 | 31 | 42 | 23 |
| U8 | 14 | 22 | 12 | 22 | 27 | 15 | 32 | 36 | 26 |
| U9 | 11 | 16 | 8 | 23 | 22 | 16 | 29 | 30 | 25 |
| U10 | 16 | 23 | 12 | 24 | 32 | 22 | 30 | 34 | 24 |
| U11 | 10 | 19 | 9 | 18 | 22 | 19 | 25 | 30 | 21 |
| U12 | 17 | 24 | 13 | 26 | 31 | 21 | 31 | 39 | 27 |
| U13 | 14 | 20 | 11 | 19 | 22 | 15 | 21 | 33 | 22 |
| U14 | 17 | 25 | 14 | 23 | 27 | 18 | 25 | 37 | 24 |
| U15 | 22 | 30 | 17 | 30 | 41 | 23 | 41 | 51 | 21 |
| Sum | 199 | 331 | 163 | 332 | 416 | 274 | 435 | 571 | 378 |
| Std | 3.49 | 3.95 | 2.53 | 4.35 | 5.54 | 4.13 | 6.86 | 6.56 | 4.75 |

**A 4: Raw data of time spent to answer questions in Control group**

| User | Level complexity | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Easy task | | | Moderate task | | | Complexity task | | |
| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 |
| U1 | 20 | 15 | 20 | 40 | 45 | 35 | 50 | 1:3 | 1:5 |
| U2 | 30 | 25 | 13 | 40 | 37 | 55 | 1:8 | 1:6 | 1:2 |
| U3 | 33 | 19 | 15 | 40 | 43 | 44 | 1:2 | 1:3 | 1:1 |
| U4 | 24 | 14 | 22 | 35 | 40 | 37 | 1: | 1:5 | 1:2 |
| U5 | 10 | 27 | 16 | 50 | 43 | 31 | 1:3 | 1:10 | 55 |
| U6 | 22 | 17 | 13 | 48 | 38 | 36 | 1: 10 | 55 | 1:1 |
| U7 | 15 | 10 | 14 | 47 | 27 | 30 | 1: 2 | 35 | 40 |
| U8 | 24 | 35 | 33 | 50 | 34 | 58 | 1:3 | 1: | 1:7 |
| U9 | 13 | 17 | 22 | 40 | 35 | 41 | 45 | 1:2 | 1: |
| U10 | 13 | 16 | 21 | 26 | 36 | 31 | 42 | 40 | 1: |
| U11 | 23 | 14 | 21 | 40 | 34 | 42 | 45 | 1:2 | 1:1 |
| U12 | 21 | 21 | 11 | 36 | 30 | 34 | 42 | 41 | 1:1 |
| U13 | 11 | 14 | 15 | 35 | 40 | 38 | 52 | 40 | 50 |
| U14 | 26 | 13 | 17 | 44 | 38 | 40 | 1:3 | 1:2 | 1: |
| U15 | 27 | 35 | 18 | 40 | 27 | 35 | 53 | 1:3 | 43 |
| Sum | 312 | 219 | 271 | 611 | 547 | 587 | 789 | 847 | 868 |
| Std | 7.04 | 7.73 | 5.45 | 6.45 | 5.50 | 8.17 | 9.52 | 11.41 | 7.70 |

**A5: Raw data of answering questions correctness in Experimental group**

**1: Correct answer, 0: Incorrect answer**

| User | Level complexity | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Easy task | | | Moderate task | | | Complexity task | | | |
| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Score |
| U1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 6 |
| U2 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 5 |
| U3 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 7 |
| U4 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 6 |
| U5 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 5 |
| U6 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 7 |
| U7 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 6 |
| U8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 |
| U9 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 5 |
| U10 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 8 |
| U11 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 8 |
| U12 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 6 |
| U13 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 |
| U14 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 8 |
| U15 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 7 |

**A6: Raw data of answering questions correctness in Control group**

**1: Correct answer, 0: Incorrect answer**

| User | Level complexity | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Easy task | | | Moderate task | | | Complexity task | | | |
| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Score |
| U1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 5 |
| U2 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 3 |
| U3 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 4 |
| U4 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 3 |
| U5 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 5 |
| U6 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 5 |
| U7 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 6 |
| U8 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 3 |
| U9 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 6 |
| U10 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 5 |
| U11 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 4 |
| U12 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 4 |
| U13 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 |
| U14 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 3 |
| U15 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 4 |

**A7: Raw data of user's response to satisfaction questionnaire**

| User | Experimental group | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 |
| 1 | 4 | 1 | 5 | 1 | 5 | 3 | 5 | 3 | 5 | 4 |
| 2 | 3 | 3 | 3 | 2 | 4 | 1 | 4 | 1 | 3 | 1 |
| 3 | 3 | 2 | 5 | 1 | 4 | 2 | 2 | 2 | 5 | 2 |
| 4 | 3 | 1 | 3 | 2 | 3 | 1 | 5 | 4 | 5 | 2 |
| 5 | 3 | 3 | 4 | 3 | 4 | 3 | 3 | 3 | 3 | 3 |
| 6 | 5 | 3 | 4 | 3 | 4 | 2 | 5 | 2 | 4 | 5 |
| 7 | 5 | 5 | 4 | 3 | 3 | 3 | 3 | 1 | 3 | 5 |
| 8 | 5 | 3 | 3 | 1 | 4 | 3 | 3 | 3 | 5 | 1 |
| 9 | 3 | 3 | 3 | 3 | 5 | 2 | 5 | 1 | 5 | 3 |
| 10 | 5 | 3 | 3 | 3 | 5 | 1 | 3 | 4 | 3 | 5 |
| 11 | 4 | 2 | 4 | 1 | 3 | 2 | 5 | 2 | 4 | 2 |
| 12 | 3 | 3 | 4 | 3 | 1 | 3 | 5 | 1 | 3 | 5 |
| 13 | 5 | 1 | 5 | 2 | 5 | 2 | 3 | 3 | 4 | 2 |
| 14 | 3 | 3 | 5 | 1 | 3 | 3 | 5 | 2 | 3 | 5 |
| 15 | 5 | 2 | 4 | 1 | 5 | 2 | 3 | 5 | 5 | 1 |
| **Score** | 98.3% | 63.2% | 98% | 50% | 95% | 58% | 98.3% | 61.6% | 100% | 76.5% |

| User | Control group | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 |
| 1 | 3 | 3 | 3 | 1 | 3 | 1 | 2 | 3 | 1 | 4 |
| 2 | 3 | 3 | 3 | 2 | 4 | 3 | 4 | 3 | 3 | 1 |
| 3 | 3 | 2 | 4 | 4 | 2 | 2 | 2 | 2 | 2 | 2 |
| 4 | 3 | 5 | 3 | 2 | 3 | 1 | 4 | 1 | 3 | 1 |
| 5 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 6 | 4 | 5 | 3 | 1 | 3 | 4 | 4 | 2 | 4 | 1 |
| 7 | 3 | 5 | 1 | 3 | 3 | 3 | 3 | 1 | 3 | 3 |
| 8 | 4 | 3 | 3 | 1 | 3 | 1 | 3 | 3 | 1 | 1 |
| 9 | 3 | 3 | 2 | 4 | 3 | 3 | 2 | 3 | 3 | 1 |
| 10 | 3 | 4 | 3 | 3 | 2 | 3 | 3 | 4 | 3 | 3 |
| 11 | 4 | 5 | 4 | 4 | 3 | 2 | 2 | 2 | 2 | 2 |
| 12 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 5 |
| 13 | 4 | 1 | 2 | 4 | 3 | 2 | 3 | 1 | 4 | 2 |
| 14 | 3 | 5 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 |
| 15 | 3 | 2 | 4 | 3 | 3 | 1 | 3 | 2 | 1 | 3 |
| **Score** | 81.6% | 85.0% | 75.00% | 68.3% | 73.2% | 55% | 71.6% | 60% | 65% | 58.2% |

## Appendix B– Experimental Phase II: The Role of Avatar, Record speech, Description Text with Images in E-assessment: A Three Condition Approach

### A.1.1: Experimental Group.

I am happy to introduce myself to you as one of the postgraduate research students in the School of Technology at the University of De Montfort t. I am at present investigating the use of multimodal metaphors in e-Assessment interfaces, and I would like to find your views about the use of multimodal metaphors such as: Avatar, also visual text and images. Please follow the next procedure:

 Answer the pre-experiment questions.

 Read every task carefully.

 Execute the task.

 In conclusion of the task, answer the essential link questions.

 After that completion of every task, answer the satisfaction questions.

Please answer every part of the question as honestly as possible. It would be appreciative if you could fill in the next questionnaire sincerely and give your views. Thank you and me greatly your participation.

**Part 1 Pre- experiment**

1- What is your age?

 18 - 24.  25 - 34.  35 - 44.  45 - 54.  55 +.

2- What is your gender?

 Male  Female

3- What is your education stage?

 High School (College (undergraduate (Master Degree (Doctorate Degree  Other Area of

Study: ………………………………………………………..

4- How usually do you use the computer (rate) per week?

 1-5 hours  6-10 hours  More than 10 hours

5- Did you perform the use of any e-learning web sites or software?

 Yes  No

6- Do you have knowledge about Multimodal ?

 Yes  No

7- What are the major reasons you use the Internet?

 Surveying  Email  Education

8- Do you think the exam in e-assessment is a good way to enhance E-learning?

 Yes  No

**“Input interface”**

**Task 1:**

- Please open program www.e-assessment.com
- Enter your name and number in the box
- Enter login in the task are coming in the first page
- Choose assessment 2

- Task 2: To achieve this task perform the following requirements and answer the questions:

- Move the mouse cursor over the "question"
- Move the mouse cursor over the multiple-choice of - the answer
- Move the mouse arrow over the answer and press now on the right answer multiple-choice of please the following question, show the images and lesson to record it might be help you to get a right answer

1- The largest coffee growing country in the world is :

- Somalia
- Yemen
- Brazil
- Move the mouse cursor over the " Next to page ".

**Task 3**: To achieve this task, perform the following requirements and answer the questions :

- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer multiple-choice of please the following question, show the images and lesson to record it might be help you to get a right answer.

2- **Which of the following is not a method of accessing the web**:

- ISDN
- Modem
- CPU

Move the mouse cursor over the " Next to page ".

**Task 4**: To achieve this task, perform the following requirements and answer the questions :

- Move the mouse cursor over the "question".
- Move the mouse cursor over the True or False of the answer.
- Move the mouse over the arrow the answer and press now on the right answer please the following question, show the image and lesson to record it might be help you to get a right answer.

3- Who invented the internet Tim Berners (     ).

- Move the mouse cursor over the " Next to page ".

**Task 5:** To achieve this task, perform the following requirements and answer the questions :

- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer, please the following question, show the visual text and images it might be help you to get a right answer.

4- Which Vitamin helps in the absorption of Calcium:

- Vitamin D
- Vitamin C
- Vitamin A

- Move the mouse cursor over the " Next to page "

**Task 6:** To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer, please the following questions, but before press the answer shows the visual text and images might be help you to get a right answer.

**5-Which is the smallest ocean in the world**

- Atlantic Ocean
- Arctic Ocean
- Indian Ocean

Move the mouse cursor over the " Next to page ".

**Task 7:** To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the of the answer.

- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer, please the following questions, but before press the answer shows the visual text and images might be help you to get a right answer.

6- The human body is made up of several chemical elements; the element present in the highest proportion (65%) in the body is:

- Carbon
- Nitrogen
- Hydrogen
- Move the mouse cursor over the " Next to page ".

**Task 8**: To achieve this task perform the following requirements and answer the questions:

- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer.
- Move the mouse over the arrow the answer and press now on the right answer, please the following questions press on the Avatar and show images it might be help you to get a right answer

7- Which of the following keys of a personal computer is not available in the keyboard of traditional typewriters

- Tab
- Spacebar
- Enter

- Move the mouse cursor over the " Next to page ".

**Task 9:** To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the "question".
- Move the mouse cursor over the True or Foals of - the answer.

• Move the mouse over the arrow the answer and press now on the right answer T or F please the following questions, but before press the answer press on the Avatar might be help you to get a right answer.

8- Electric current is measured by the voltmeter ()

- Move the mouse cursor over the " Next to page "

**Task 10:** To achieve this task, perform the following requirements and answer the questions:

- Move the mouse cursor over the "question".
- Move the mouse cursor over the multiple-choice of - the answer
- Move the mouse over the arrow the answer and press now on the right answer the multiple-choice of please the following questions, but before press the answer press on the Avatar might be help you to get a right answer

9- How many hearts does an octopus have?

- One
- Two
- Three

- Move the mouse cursor over the “Finish “to show your time and your score and click to logout.

## Satisfaction Questionnaire

For each statement below, please express your view by placing a circle in the appropriate column.

**5**= **(SA)** Strongly Agree. **4**= **(A)** Agree. **3**= **(N)** Natural. **2**= **(D)** Disagree. **1**= **(SD)** Strongly Disagree.

| **Statement** | | **(SA)** | **(A)** | **(N)** | **(D)** | **(SD)** |
|---|---|---|---|---|---|---|
| S1 | I think that I would like to use this system frequently | 5 | 4 | 3 | 2 | 1 |
| S2 | I found the system unnecessarily complex | 5 | 4 | 3 | 2 | 1 |
| S3 | I thought the system was easy to use | 5 | 4 | 3 | 2 | 1 |
| S4 | I think that I would need the support of a technical person to be able to use this system | 5 | 4 | 3 | 2 | 1 |
| S5 | I found the various functions in this system were well integrated | 5 | 4 | 3 | 2 | 1 |
| S6 | I thought there was too much inconsistency in this system | 5 | 4 | 3 | 2 | 1 |
| S7 | I would imagine that most people would learn to use this system very quickly | 5 | 4 | 3 | 2 | 1 |
| S8 | I found the system very cumbersome to use | 5 | 4 | 3 | 2 | 1 |
| S9 | I felt very confident using the system | 5 | 4 | 3 | 2 | 1 |
| S10 | I needed to learn a lot of things before I could get going with this system | 5 | 4 | 3 | 2 | 1 |

If you have any comments or suggestions, can you write please?

**B2: User Profile Data**

| | | Experiment group | Control Group |
|---|---|---|---|
| Age | 18- 24 | 80% | 60% |
| | 25- 30 | 25% | 40% |
| | 31- 40 | | |
| Gender | Male | 40% | 60% |
| | Female | 60% | 40% |
| Education level | Undergraduate level | 73% | 67% |
| | Postgraduate level | 26.7% | 33% |
| Use Of computer | 1 – 5 hours | 26.7% | 13.3% |
| | 6 – 10 hours | 26.7% | 40% |
| | More than 10 hours | 46.7% | 46.7% |
| E-learning experience | Yes | 53.3% | 66.7% |
| | No | 46.7% | 33.3% |
| Multimodal Knowledge | Yes | 66.7% | 53.3% |
| | No | 33.3% | 46.7% |
| Internet use reason | Surveying | 60% | 53.3% |
| | Email | 13.3% | 20% |
| | Education | 26.7% | 26.7% |
| E-assessment enhance E-learning | Yes | 80% | 60% |
| | No | 40% | 40% |

**B3: Time spent to answer Assessment questions for Assessment types**

| User | Different multimodal | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Record speech with images task RI | | | description text with images task DI | | | Avatar with images task AI | | |
| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 |
| U1 | 20 | 20 | 31 | 16 | 29 | 30 | 10 | 15 | 23 |
| U2 | 19 | 24 | 31 | 19 | 31 | 41 | 10 | 14 | 20 |
| U3 | 16 | 19 | 29 | 20 | 21 | 40 | 10 | 18 | 23 |
| U4 | 20 | 22 | 28 | 18 | 27 | 33 | 13 | 19 | 18 |
| U5 | 18 | 32 | 29 | 21 | 23 | 30 | 9 | 15 | 19 |
| U6 | 10 | 25 | 19 | 16 | 20 | 41 | 11 | 15 | 21 |
| U7 | 20 | 27 | 27 | 22 | 24 | 33 | 10 | 16 | 20 |
| U8 | 18 | 18 | 28 | 18 | 23 | 41 | 9 | 13 | 24 |
| U9 | 20 | 19 | 31 | 17 | 30 | 31 | 15 | 18 | 21 |
| U10 | 14 | 20 | 27 | 14 | 30 | 40 | 10 | 13 | 22 |
| U11 | 21 | 19 | 31 | 20 | 30 | 43 | 12 | 17 | 20 |
| U12 | 15 | 22 | 27 | 19 | 27 | 30 | 17 | 13 | 23 |
| U13 | 12 | 26 | 38 | 18 | 25 | 30 | 9 | 10 | 25 |
| U14 | 21 | 24 | 21 | 21 | 23 | 34 | 12 | 13 | 18 |
| U15 | 19 | 26 | 35 | 19 | 31 | 31 | 13 | 15 | 17 |
| U16 | 20 | 20 | 31 | 22 | 29 | 36 | 10 | 15 | 23 |
| U17 | 20 | 24 | 31 | 20 | 33 | 41 | 8 | 18 | 20 |
| U18 | 16 | 19 | 29 | 15 | 32 | 30 | 16 | 18 | 22 |
| U19 | 20 | 19 | 28 | 17 | 23 | 42 | 8 | 19 | 18 |
| U20 | 18 | 29 | 29 | 21 | 23 | 33 | 13 | 15 | 24 |
| U21 | 18 | 18 | 35 | 20 | 26 | 41 | 11 | 15 | 21 |
| U22 | 19 | 32 | 27 | 16 | 22 | 30 | 10 | 16 | 20 |
| U23 | 15 | 33 | 28 | 18 | 23 | 31 | 9 | 18 | 24 |
| U24 | 20 | 19 | 31 | 18 | 25 | 40 | 7 | 18 | 21 |
| U25 | 14 | 20 | 27 | 19 | 22 | 33 | 10 | 13 | 22 |
| U26 | 21 | 21 | 31 | 20 | 30 | 41 | 12 | 17 | 20 |
| U27 | 15 | 22 | 27 | 19 | 27 | 30 | 14 | 20 | 23 |
| U28 | 12 | 19 | 26 | 15 | 25 | 31 | 9 | 10 | 25 |
| U29 | 14 | 19 | 40 | 20 | 23 | 34 | 12 | 13 | 18 |
| U30 | 14 | 26 | 35 | 19 | 30 | 30 | 13 | 15 | 22 |
| SUM | 17.3 | 22.7 | 29.5 | 18.5 | 26.2 | 35.03 | 11.06 | 15.4 | 21.2 |

**B4: Raw data of time spent to watch Assessment multimodal in each condition**

| User | Different multimodal | | |
|---|---|---|---|
| | Record speech with images task | description text with images task | Avatar with images task |
| | RI | DI | AI |
| U1 | 71 | 75 | 48 |
| U2 | 74 | 91 | 44 |
| U3 | 64 | 81 | 51 |
| U4 | 70 | 78 | 50 |
| U5 | 79 | 74 | 43 |
| U6 | 54 | 77 | 47 |
| U7 | 74 | 79 | 46 |
| U8 | 64 | 82 | 46 |
| U9 | 70 | 78 | 54 |
| U10 | 61 | 84 | 45 |
| U11 | 71 | 93 | 49 |
| U12 | 64 | 76 | 53 |
| U13 | 76 | 73 | 44 |
| U14 | 66 | 78 | 43 |
| U15 | 80 | 81 | 45 |
| U16 | 71 | 87 | 48 |
| U17 | 75 | 94 | 46 |
| U18 | 64 | 77 | 56 |
| U19 | 67 | 82 | 45 |
| U20 | 76 | 77 | 52 |
| U21 | 71 | 87 | 47 |
| U22 | 78 | 68 | 46 |
| U23 | 76 | 72 | 51 |
| U24 | 70 | 83 | 46 |
| U25 | 61 | 74 | 45 |
| U26 | 73 | 91 | 49 |
| U27 | 64 | 76 | 57 |
| U28 | 57 | 71 | 44 |
| U29 | 73 | 77 | 43 |
| U30 | 75 | 79 | 50 |
| SUM | 69.6 | 79.8 | 47.7 |

**B5: Raw data of time spent to watch Assessment time complexity Task**

| User | Level complexity | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Easy task | | | Moderate task | | | Complexity task | | |
| | T1RI | T4DI | T7AI | T2RI | T5DI | T8AI | T3RI | T6DI | T9AI |
| U1 | 20 | 16 | 10 | 20 | 29 | 15 | 31 | 30 | 23 |
| U2 | 19 | 19 | 10 | 24 | 31 | 14 | 31 | 41 | 20 |
| U3 | 16 | 20 | 10 | 19 | 21 | 18 | 29 | 40 | 23 |
| U4 | 20 | 18 | 13 | 22 | 27 | 19 | 28 | 33 | 18 |
| U5 | 18 | 21 | 9 | 32 | 23 | 15 | 29 | 30 | 19 |
| U6 | 10 | 16 | 11 | 25 | 20 | 15 | 19 | 41 | 21 |
| U7 | 20 | 22 | 10 | 27 | 24 | 16 | 27 | 33 | 20 |
| U8 | 18 | 18 | 9 | 18 | 23 | 13 | 28 | 41 | 24 |
| U9 | 20 | 17 | 15 | 19 | 30 | 18 | 31 | 31 | 21 |
| U10 | 14 | 14 | 10 | 20 | 30 | 13 | 27 | 40 | 22 |
| U11 | 21 | 20 | 12 | 19 | 30 | 17 | 31 | 43 | 20 |
| U12 | 15 | 19 | 17 | 22 | 27 | 13 | 27 | 30 | 23 |
| U13 | 12 | 18 | 9 | 26 | 25 | 10 | 38 | 30 | 25 |
| U14 | 21 | 21 | 12 | 24 | 23 | 13 | 21 | 34 | 18 |
| U15 | 19 | 19 | 13 | 26 | 31 | 15 | 35 | 31 | 17 |
| U16 | 20 | 22 | 10 | 20 | 29 | 15 | 31 | 36 | 23 |
| U17 | 20 | 20 | 8 | 24 | 33 | 18 | 31 | 41 | 20 |
| U18 | 16 | 15 | 16 | 19 | 32 | 18 | 29 | 30 | 22 |
| U19 | 20 | 17 | 8 | 19 | 23 | 19 | 28 | 42 | 18 |
| U20 | 18 | 21 | 13 | 29 | 23 | 15 | 29 | 33 | 24 |
| U21 | 18 | 20 | 11 | 18 | 26 | 15 | 35 | 41 | 21 |
| U22 | 19 | 16 | 10 | 32 | 22 | 16 | 27 | 30 | 20 |
| U23 | 15 | 18 | 9 | 33 | 23 | 18 | 28 | 31 | 24 |
| U24 | 20 | 18 | 7 | 19 | 25 | 18 | 31 | 40 | 21 |
| U25 | 14 | 19 | 10 | 20 | 22 | 13 | 27 | 33 | 22 |
| U26 | 21 | 20 | 12 | 21 | 30 | 17 | 31 | 41 | 20 |
| U27 | 15 | 19 | 14 | 22 | 27 | 20 | 27 | 30 | 23 |
| U28 | 12 | 15 | 9 | 19 | 25 | 10 | 26 | 31 | 25 |
| U29 | 14 | 20 | 12 | 19 | 23 | 13 | 40 | 34 | 18 |
| U30 | 14 | 19 | 13 | 26 | 30 | 15 | 35 | 30 | 22 |
| Average | 17.3 | 18.5 | 11.06 | 22.7 | 26.2 | 15.4 | 29.5 | 35.03 | 21.2 |

**B6: Raw data of answering questions correctness and level complexity of each user Record speech with images Condition (RI) description text with images Condition (DI) Avatar with images Condition (AI) level complexity ( E ) Easy ( M ) Moderate ( D ) Difficult**

| User | Level complexity | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | RI condition | | | DI condition | | | AI condition | | | |
| | QI E | Q2 M | Q3 D | Q4 E | Q5 M | Q6 D | Q7 E | Q8 M | Q9 D | Score |
| U1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 4 |
| U2 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 6 |
| U3 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 7 |
| U4 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 4 |
| U5 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 5 |
| U6 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 7 |
| U7 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 5 |
| U8 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 8 |
| U9 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 7 |
| U10 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 6 |
| U11 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 7 |
| U12 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 2 |
| U13 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 6 |
| U14 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 7 |
| U15 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 6 |
| U16 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 6 |
| U17 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 6 |
| U18 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 6 |
| U19 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 5 |
| U20 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 6 |
| U21 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 4 |
| U22 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 2 |
| U23 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 6 |
| U24 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 6 |
| U25 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 6 |
| U26 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 5 |
| U27 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 5 |
| U28 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 7 |
| U29 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 5 |
| U30 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 7 |
| SUM | 24 | 21 | 17 | 19 | 14 | 11 | 28 | 25 | 19 | |
| Average | 62% | | | 44% | | | 72% | | | 6 |

**B7: Raw data of satisfaction questionnaire score of each satisfaction statement of Record and images condition**

| Record and images condition | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| USER | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | Score |
| U1 | 1 | 3 | 1 | 3 | 3 | 3 | 1 | 3 | 4 | 3 | 62.5% |
| U2 | 1 | 4 | 1 | 1 | 1 | 1 | 4 | 2 | 4 | 4 | 60% |
| U3 | 4 | 4 | 1 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U4 | 4 | 1 | 3 | 1 | 2 | 2 | 3 | 2 | 4 | 1 | 57.5% |
| U5 | 2 | 2 | 2 | 2 | 3 | 3 | 4 | 4 | 2 | 2 | 65% |
| U6 | 1 | 4 | 3 | 3 | 1 | 1 | 2 | 1 | 4 | 4 | 60% |
| U7 | 2 | 2 | 1 | 1 | 3 | 3 | 2 | 3 | 3 | 2 | 100% |
| U8 | 3 | 1 | 3 | 2 | 2 | 2 | 3 | 1 | 4 | 1 | 100% |
| U9 | 1 | 4 | 2 | 1 | 4 | 2 | 1 | 2 | 1 | 3 | 52.5% |
| U10 | 3 | 1 | 2 | 2 | 3 | 3 | 2 | 2 | 4 | 1 | 57.5% |
| U11 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U12 | 1 | 3 | 4 | 2 | 4 | 2 | 1 | 3 | 2 | 1 | 57.50% |
| U13 | 1 | 3 | 1 | 1 | 3 | 1 | 1 | 2 | 4 | 3 | 50% |
| U14 | 3 | 1 | 3 | 3 | 4 | 1 | 3 | 1 | 3 | 2 | 60% |
| U15 | 2 | 4 | 1 | 4 | 1 | 1 | 1 | 4 | 4 | 1 | 57.50% |
| U16 | 2 | 2 | 2 | 1 | 4 | 3 | 1 | 1 | 3 | 2 | 100% |
| U17 | 3 | 1 | 4 | 4 | 4 | 4 | 4 | 4 | 2 | 2 | 80% |
| U18 | 1 | 3 | 2 | 2 | 3 | 3 | 4 | 4 | 2 | 3 | 67.5% |
| U19 | 1 | 4 | 1 | 1 | 3 | 3 | 1 | 2 | 4 | 4 | 60% |
| U20 | 1 | 3 | 4 | 1 | 4 | 4 | 2 | 1 | 2 | 3 | 62.5% |
| U21 | 3 | 1 | 3 | 1 | 2 | 4 | 1 | 4 | 1 | 1 | 52.5% |
| U22 | 3 | 3 | 1 | 3 | 2 | 4 | 4 | 2 | 3 | 1 | 65% |
| U23 | 2 | 2 | 2 | 1 | 4 | 1 | 3 | 1 | 3 | 2 | 52.5% |
| U24 | 1 | 4 | 3 | 3 | 4 | 4 | 1 | 3 | 1 | 4 | 70% |
| U25 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U26 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U27 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U28 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U29 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U30 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| **Average** | | | | | | | | | | | 75% |

**B8: Raw data of satisfaction questionnaire score of each satisfaction statement of description text and images condition**

| description text and images condition | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| USER | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | Score |
| U1 | 3 | 1 | 3 | 1 | 3 | 3 | 2 | 4 | 3 | 4 | 67.5% |
| U2 | 2 | 4 | 1 | 3 | 4 | 1 | 1 | 1 | 3 | 4 | 60% |
| U3 | 1 | 2 | 1 | 3 | 1 | 4 | 4 | 1 | 1 | 2 | 50% |
| U4 | 3 | 1 | 3 | 1 | 2 | 2 | 3 | 2 | 3 | 1 | 52.5% |
| U5 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 1 | 2 | 1 | 52.5% |
| U6 | 2 | 4 | 3 | 3 | 1 | 1 | 2 | 1 | 2 | 4 | 57.5% |
| U7 | 3 | 2 | 1 | 1 | 3 | 2 | 2 | 1 | 3 | 2 | 50% |
| U8 | 4 | 3 | 3 | 4 | 2 | 2 | 3 | 1 | 2 | 1 | 62.5% |
| U9 | 1 | 4 | 4 | 1 | 2 | 2 | 4 | 2 | 1 | 3 | 60% |
| U10 | 1 | 1 | 3 | 2 | 3 | 3 | 2 | 1 | 1 | 1 | 45% |
| U11 | 4 | 4 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 45% |
| U12 | 2 | 3 | 4 | 2 | 1 | 1 | 1 | 3 | 2 | 1 | 50% |
| U13 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 3 | 100% |
| U14 | 3 | 2 | 1 | 3 | 4 | 2 | 3 | 2 | 3 | 2 | 62.5% |
| U15 | 4 | 4 | 3 | 2 | 1 | 1 | 4 | 1 | 4 | 1 | 62.5% |
| U16 | 3 | 1 | 2 | 1 | 1 | 3 | 4 | 1 | 3 | 2 | 52.5% |
| U17 | 1 | 2 | 4 | 4 | 4 | 1 | 2 | 2 | 4 | 2 | 65% |
| U18 | 2 | 3 | 2 | 2 | 3 | 3 | 4 | 4 | 2 | 3 | 60% |
| U19 | 1 | 1 | 1 | 4 | 1 | 2 | 1 | 4 | 1 | 1 | 100% |
| U20 | 1 | 3 | 1 | 1 | 4 | 4 | 2 | 2 | 2 | 3 | 57.5% |
| U21 | 1 | 2 | 3 | 3 | 2 | 3 | 1 | 2 | 1 | 1 | 47.5% |
| U22 | 2 | 1 | 1 | 3 | 2 | 2 | 4 | 2 | 3 | 1 | 52.5% |
| U23 | 2 | 2 | 2 | 1 | 3 | 2 | 3 | 1 | 3 | 2 | 52.5% |
| U24 | 1 | 4 | 3 | 3 | 4 | 1 | 4 | 3 | 1 | 4 | 70% |
| U25 | 1 | 2 | 1 | 1 | 4 | 2 | 4 | 2 | 4 | 4 | 62.5% |
| U26 | 2 | 2 | 2 | 3 | 1 | 3 | 1 | 3 | 2 | 1 | 50% |
| U27 | 2 | 3 | 1 | 2 | 2 | 3 | 3 | 3 | 2 | 1 | 57.5% |
| U28 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | %100 |
| U29 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | %100 |
| U30 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | %100 |
| **Average** | | | | | | | | | | | 63.5% |

**B9: Raw data of satisfaction questionnaire score of each satisfaction statement of Avatar and images condition**

| Avatar and images condition | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| USER | S 1 | S 2 | S 3 | S 4 | S 5 | S 6 | S 7 | S 8 | S 9 | S10 | Score |
| U1 | 3 | 4 | 4 | 2 | 4 | 4 | 4 | 3 | 3 | 4 | 87.5% |
| U2 | 4 | 2 | 1 | 4 | 3 | 1 | 3 | 4 | 4 | 4 | 75% |
| U3 | 3 | 4 | 4 | 3 | 4 | 3 | 2 | 1 | 3 | 2 | 72.5% |
| U4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U5 | 1 | 4 | 4 | 4 | 4 | 1 | 4 | 4 | 2 | 2 | 75% |
| U6 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U7 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U8 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U9 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U10 | 4 | 1 | 2 | 4 | 3 | 3 | 2 | 4 | 4 | 1 | 70% |
| U11 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U12 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U13 | 4 | 3 | 4 | 1 | 3 | 4 | 3 | 4 | 4 | 3 | 82.5% |
| U14 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U15 | 4 | 4 | 4 | 4 | 1 | 2 | 4 | 2 | 2 | 1 | 70% |
| U16 | 3 | 3 | 2 | 2 | 4 | 3 | 2 | 1 | 3 | 4 | 67.5% |
| U17 | 2 | 4 | 3 | 4 | 3 | 4 | 4 | 4 | 4 | 2 | 85% |
| U18 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U19 | 4 | 4 | 1 | 4 | 3 | 3 | 4 | 2 | 4 | 4 | 82.5% |
| U20 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U21 | 4 | 4 | 3 | 3 | 4 | 3 | 4 | 4 | 4 | 4 | 92.5% |
| U22 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 1 | 100% |
| U23 | 4 | 2 | 2 | 4 | 4 | 1 | 3 | 2 | 3 | 2 | 67.5% |
| U24 | 4 | 4 | 4 | 2 | 3 | 4 | 4 | 3 | 1 | 4 | 82.5% |
| U25 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 1 | 4 | 4 | 100% |
| U26 | 2 | 2 | 2 | 3 | 4 | 4 | 3 | 3 | 2 | 2 | 67.5% |
| U27 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100% |
| U28 | 4 | 4 | 4 | 4 | 3 | 2 | 2 | 4 | 2 | 4 | 82.5% |
| U29 | 4 | 4 | 4 | 2 | 4 | 4 | 4 | 4 | 3 | 1 | 85% |
| U30 | 4 | 4 | 4 | 4 | 3 | 3 | 4 | 4 | 4 | 4 | 95% |
| Average | | | | | | | | | | | 88% |

## Appendix C: Experimental Phase III: The Role of Expressive Body Gesture and Earcons and Auditory Icons in E-assessment Interfaces

I am pleased to present myself to you as one of the postgraduate research students in the Faculty of Technology at the De Montfort University. I am currently investigating the use of multimodal metaphors in e-assessment interfaces, and I would like to obtain your views regarding the use of auditory non-speech metaphors in e-assessment interface.

Please complete the following procedure:

 Answer the pre-experiment questions.

 Execute the tasks.

 Then answer recall and recognitions questions.

 Evaluate each auditory non-speech metaphors used.

 After that, answer the satisfaction questionnaire.

Please complete all the requirements as honestly as possible. It would be grateful if you could fill in the following questionnaire sincerely and express your views. Your privacy is guaranteed as you will not be mentioned in any part of the study. Thank you very much, and I highly appreciate your contribution.

**Part 1 Pre- experiment**

1- What is your age?

 18 - 24.  25 - 34.  35 - 44.  45 - 54.  55 +.

2- What is you gender?

 Male  Female

3- What is your education stage?

 High School  College  Under-graduate  Master Degree  Doctorate Degree  Other Area of Study: ………………………………………………………..

4- How usually do you use the computer (rate) per week?

 1-5 hours  6-10 hours  More than 10 hours

5- Did you perform the use of any e-learning web sites or software?

 Yes  No

6- Do you have knowledge about Multimodal ?

 Yes  No

7- What are the major reasons you use the Internet?

 Surveying  Email  Education

8- Do you think the exam in E-assessment is good way to enhance E-learning?

 Yes  No

**Tasks:**

**Achievement stage:**

***Part 1:***

In this experiment you will see and listen to the assessment that provide via tutor. It should be noted there are six dissimilar types of assistance content you see or listen. These assistance types are as next Error, Comments, involving Thinking, Explain question, more Suggestions and Mark. First, you need to press on Error button in the top of the condition (at this moment you will hear auditory non-speech) then you will see the tutor talk on the left of the condition. At same time you will see the text about Error on the center of the condition.

Then you need to press on Comments button on the in the top of the condition (at this moment you will hear auditory non-speech) after that you will see the tutor talk on the left of the condition. At same time you will see the text about Comments on the center of the condition. It is requested to focus on what is provided because you will be asked some questions about that. Then press on Questions button at the bottom of the condition, in this page you need to answer these questions select correct answer. These questions about what are provide on the previous condition regarding assessment provided.

1- What type of Error the instructor mentioned to?

……………………………………………………………………………

2- The part that supposed out well and determined strength and weakness of the method is: (select one)

a- Preface b- key body c-literature review d- illation.

After completing correct answering questions press on following button. In this question you ask to evaluation the technique that presented assessment. Thus, rate each statement via tick on the appropriate to you.

***Part 2:***

Thereafter, press on following button. Secondly, you need to press on Involve Thinking button in the highest of the interface (at the same this moment you will get auditory non-speech) then you will listen to the teacher speaking around this type of assistance. At same time you will get the text about Involve Thinking on the central of the interface. Then you need to press on Explain question button on the in the highest of the interface (at the same this moment you will listen to the auditory non-speech) then you will see the teacher speaking around this type of assistance. At same time you will see the text around Explain question on the central of the interface. It is order to concentrate on what is offered because you will be requested some questions about that. Then, press on Questions button at the bottom of the interface, in this sheet you need to answer these questions also writes answer or select correct answer. These questions around what are offered on the previous interface regarding assessment offered.

1- To which part the instructor involved user thinking in conclusion?

……………………………………………………………………………

2- What is the user requested to do in this part: (Select one)

a- Investigate using new method b- Get advantages with the new methodology

c- Get disadvantages with the new methodology

After completing correct answering questions press on following button. In this question you ask to evaluation the technique that presented assessment. Thus, rate each statement via tick on the appropriate to you.

***Part 3:***

Thereafter, press on following button .Thirdly, you need to press on More Suggestions button in the top of the interface (at the same this moment you will get auditory non-speech) then you will see the teacher conversation on the left of the interface. At same time you will get the text about More Suggestions on the central of the interface. Then you need to press on Marks button on the in the top of the interface (at this moment you will listen to the auditory non-speech) then you will get the teacher conversation on the left of the interface. At same time you will see the text around Marks on the central of the interface. It is order to concentrate on what is offered because you will be requested some questions about that. Then, press on Questions button at the bottom of the interface, in this sheet you need to answer these questions also writes answer or select correct answer. These questions around what are offered on the previous interface regarding assessment offered.

1- What is suggested to user to record for each task?

……………………………………………………………...

2- Which mark is assumed to Presentation part? (Select one)

a- 3/5 b- 5/10 c- 4/5 d- 40/50

**Involving Test:**

Part 1:

The subsequent six assistance types is going to be submission, point out where sound used to indicate each of the next:

- Error.
- Comment.
- Involve Thinking.
- Explain Question.
- More Suggestion.
- Mark.
- Assistance high Important.
- Assistance Medium Important.
- Assistance, Low Important.

**Part 2:**

In this experiment you will hear two sounds for each of assessment types and assistance ability level.

| Assistance types/stage | Non-speech sounds |
|---|---|
| Error | |
| Comment | |
| Involve Thinking | |
| Explain Question | |
| More Suggestion | |
| Mark | |
| Assessment high Important | |
| Assessment Medium Important | |
| Assessment Low Important | |

Select the correct sound (write number 1, 2,…. In box).

**Valuation of Non-speech sounds:**

How did you discovery the use of the further Non-speech sounds in the tested e-assessment interface? (Tick in appropriate place).

| Agree | Feeling | Disagree |
|---|---|---|
| | Irritation | |
| | frustration | |
| | Usefulness | |
| | Concentration | |

## Satisfaction:

For each statement below, please express your view by placing a circle in the appropriate column.

5= (SA) Strongly Agree. 4= (A) Agree. 3= (N) Neutral. 2= (D) Disagree. 1= (SD) Strongly Disagree

| | Statement | SA | A | N | D | S D |
|---|---|---|---|---|---|---|
| S 1 | I think that I would like to use this system frequently | 5 | 4 | 3 | 2 | 1 |
| S 2 | I found the system unnecessarily complex | 5 | 4 | 3 | 2 | 1 |
| S 3 | I thought the system was easy to use | 5 | 4 | 3 | 2 | 1 |
| S 4 | There have been time while interacting with the method where I felt worried | 5 | 4 | 3 | 2 | 1 |
| S 5 | I found the various functions in this system were well integrated | 5 | 4 | 3 | 2 | 1 |
| S 6 | I thought there was too much inconsistency in this system | 5 | 4 | 3 | 2 | 1 |
| S 7 | I would imagine that most people would learn to use this system very quickly | 5 | 4 | 3 | 2 | 1 |
| S 8 | I found the system very cumbersome to use | 5 | 4 | 3 | 2 | 1 |
| S 9 | I felt very confident using the system | 5 | 4 | 3 | 2 | 1 |
| S10 | I needed to learn a lot of things before I could get going with this system | 5 | 4 | 3 | 2 | 1 |

**C2: User Profile Data:**

| | | |
|---|---|---|
| **Age** | 18- 24 | 43% |
| | 25- 30 | 46% |
| | 31- 40 | 10% |
| **Gender** | Male | 60% |
| | Female | 40% |
| **Education level** | Undergraduate level | 56% |
| | postgraduate level | 43% |
| **Use of computer** | 1 – 5 hours | 13% |
| | 6 – 10 hours | 26 % |
| | More than 10 hours | 60 % |
| **E-learning experience** | Yes | 76 % |
| | No | 23 % |
| **Multimodal Knowledge** | Yes | 76 % |
| | No | 23 % |
| **Internet use reason** | Surveying | 66% |
| | Email | 10% |
| | Education | 23% |
| **e-assessment enhance e-learning** | Yes | 26% |
| | No | 73% |

**C3: Raw data of Correctness Answer for Achievement Level of Assistance Types and Assessment Questions Type**

| | Assistance Type | | | | | | Assessment Questions Type | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Recall | | | Recognition | | | |
| User | Error | Comment | Involving Thinking | Explain Questions | More Suggestio | Mark | Q1 | Q2 | Q3 | Q1 | Q2 | Q3 | Score |
| 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 2 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| 3 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 4 |
| 4 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 5 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 5 |
| 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 4 |
| 7 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| 8 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 4 |
| 9 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 4 |
| 10 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 5 |
| 11 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 12 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 4 |
| 13 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 14 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 5 |
| 15 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 5 |
| 16 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| 17 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 18 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 19 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 5 |
| 20 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 4 |
| 21 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 4 |
| 22 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| 23 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 5 |
| 24 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| 25 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 4 |
| 26 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| 27 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 28 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 29 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 30 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 5 |

## C4: Raw Data of Correctness Answers for Involving Test 1

| User ID | Assistance Type | | | | | | Assistance Important level | | |
|---|---|---|---|---|---|---|---|---|---|
| | Error | Comment | Involving Thinking | Explain Questions | More Suggestion | Mark | assistance high ability | assistance Medium ability | assistance Low ability |
| 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| 4 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| 5 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 6 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| 7 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| 9 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 |
| 10 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 11 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 12 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| 13 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 14 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 15 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 16 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| 17 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| 18 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 19 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 20 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| 21 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 |
| 22 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| 23 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 24 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |
| 25 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 26 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| 27 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 28 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 |
| 29 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 30 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |

## C5: Raw data of correctness answers for Involving test 2

| User ID | Assistance Type | | | | | | Assistance ability level | | |
|---|---|---|---|---|---|---|---|---|---|
| | Error | Comment | Involving Thinking | Explain Question | More Suggesting | Mark | assistance high ability | assistance Medium ability | assistance Low ability |
| 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 4 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 6 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| 7 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 9 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| 10 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 11 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| 12 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| 13 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 14 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 15 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 16 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| 17 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| 18 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 19 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 20 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| 21 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 22 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 23 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 24 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 25 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| 26 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| 27 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 28 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| 29 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 30 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |

## C6: Raw Data of User's evaluation of Non-speech Sounds used

Agree: 1. Disagree: 0.

| User ID | Usefulness | Concentration | Irritation | frustration |
|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 0 | 0 |
| 3 | 1 | 0 | 1 | 0 |
| 4 | 1 | 1 | 0 | 1 |
| 5 | 1 | 1 | 1 | 0 |
| 6 | 1 | 1 | 1 | 1 |
| 7 | 1 | 1 | 1 | 0 |
| 8 | 1 | 1 | 1 | 1 |
| 9 | 1 | 0 | 1 | 0 |
| 10 | 0 | 1 | 1 | 1 |
| 11 | 0 | 1 | 1 | 1 |
| 12 | 1 | 1 | 1 | 0 |
| 13 | 1 | 1 | 1 | 1 |
| 14 | 1 | 1 | 1 | 1 |
| 15 | 1 | 1 | 0 | 1 |
| 16 | 1 | 0 | 1 | 0 |
| 17 | 1 | 1 | 1 | 0 |
| 18 | 1 | 1 | 0 | 0 |
| 19 | 1 | 0 | 1 | 0 |
| 20 | 1 | 1 | 0 | 1 |
| 21 | 1 | 1 | 1 | 1 |
| 22 | 0 | 1 | 0 | 0 |
| 23 | 1 | 0 | 0 | 1 |
| 24 | 1 | 1 | 1 | 0 |
| 25 | 1 | 1 | 0 | 1 |
| 26 | 1 | 1 | 0 | 1 |
| 27 | 0 | 0 | 1 | 1 |
| 28 | 1 | 0 | 1 | 0 |
| 29 | 1 | 1 | 1 | 1 |
| 30 | 1 | 1 | 1 | 0 |

### C7: Raw Data of Satisfaction Questionnaire for Each Satisfaction Statement

**5**= **(SA)** Strongly Agree. **4**= **(A)** Agree. **3**= **(N)** Natural. **2**= **(D)** Disagree. **1**= **(SD)** Strongly Disagree.

| User ID | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | Score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 4 | 3 | 3 | 3 | 4 | 4 | 3 | 3 | 3 | 4 | 85 |
| 2 | 3 | 4 | 4 | 3 | 4 | 4 | 4 | 4 | 4 | 3 | 92.5 |
| 3 | 4 | 4 | 4 | 2 | 4 | 2 | 3 | 4 | 4 | 4 | 87.5 |
| 4 | 4 | 4 | 4 | 4 | 3 | 4 | 4 | 3 | 3 | 4 | 92.5 |
| 5 | 2 | 2 | 3 | 4 | 3 | 3 | 3 | 4 | 4 | 4 | 80 |
| 6 | 4 | 2 | 3 | 1 | 4 | 4 | 3 | 4 | 3 | 4 | 80 |
| 7 | 3 | 3 | 4 | 2 | 2 | 4 | 4 | 3 | 4 | 4 | 82.5 |
| 8 | 4 | 4 | 4 | 3 | 4 | 2 | 3 | 4 | 4 | 4 | 90 |
| 9 | 3 | 4 | 4 | 3 | 2 | 4 | 4 | 2 | 3 | 3 | 80 |
| 10 | 4 | 3 | 3 | 4 | 4 | 1 | 2 | 4 | 4 | 4 | 82.5 |
| 11 | 1 | 4 | 3 | 2 | 3 | 3 | 3 | 4 | 4 | 2 | 72.5 |
| 12 | 2 | 4 | 4 | 3 | 2 | 1 | 4 | 3 | 4 | 3 | 75 |
| 13 | 3 | 3 | 4 | 2 | 4 | 3 | 3 | 4 | 3 | 4 | 82.5 |
| 14 | 4 | 2 | 4 | 3 | 4 | 3 | 2 | 4 | 4 | 4 | 85 |
| 15 | 3 | 4 | 4 | 2 | 3 | 4 | 4 | 3 | 3 | 4 | 85 |
| 16 | 3 | 4 | 4 | 3 | 4 | 2 | 4 | 4 | 4 | 4 | 90 |
| 17 | 4 | 2 | 3 | 3 | 4 | 4 | 3 | 3 | 3 | 3 | 80 |
| 18 | 3 | 4 | 4 | 4 | 4 | 3 | 4 | 3 | 4 | 4 | 92.5 |
| 19 | 4 | 1 | 4 | 3 | 3 | 4 | 4 | 3 | 4 | 3 | 82.5 |
| 20 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 4 | 4 | 3 | 77.5 |
| 21 | 2 | 4 | 4 | 2 | 4 | 3 | 3 | 4 | 4 | 4 | 85 |
| 22 | 4 | 2 | 4 | 2 | 3 | 3 | 4 | 4 | 3 | 2 | 77.5 |
| 23 | 3 | 4 | 4 | 4 | 3 | 1 | 3 | 2 | 4 | 3 | 77.5 |
| 24 | 4 | 4 | 4 | 3 | 4 | 4 | 4 | 4 | 4 | 3 | 95 |
| 25 | 5 | 1 | 5 | 1 | 4 | 2 | 2 | 1 | 3 | 2 | 65 |
| 26 | 4 | 1 | 3 | 1 | 3 | 1 | 5 | 4 | 3 | 3 | 70 |
| 27 | 3 | 4 | 4 | 2 | 4 | 1 | 3 | 1 | 4 | 4 | 75 |
| 28 | 1 | 3 | 4 | 3 | 3 | 4 | 4 | 3 | 1 | 3 | 72.5 |
| 29 | 2 | 2 | 3 | 4 | 4 | 3 | 4 | 3 | 4 | 1 | 75 |
| 30 | 4 | 3 | 5 | 3 | 2 | 3 | 5 | 2 | 2 | 1 | 75 |